(12) INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(19) World Intellectual Property Organization
International Bureau

(43) International Publication Date
25 January 2001 (25.01.2001)

PCT

(10) International Publication Number
WO 01/05970 A2

(51) International Patent Classification[7]: C12N 15/12, C07K 14/47, G01N 33/53, C12Q 1/68, A61K 38/17, C07K 16/18, A01K 67/027

(21) International Application Number: PCT/US00/19698

(22) International Filing Date: 19 July 2000 (19.07.2000)

(25) Filing Language: English

(26) Publication Language: English

(30) Priority Data:

60/144,595	19 July 1999 (19.07.1999)	US
60/150,460	23 August 1999 (23.08.1999)	US
60/159,849	15 October 1999 (15.10.1999)	US

(63) Related by continuation (CON) or continuation-in-part (CIP) to earlier applications:

US	60/144,595 (CIP)
Filed on	19 July 1999 (19.07.1999)
US	60/150,460 (CIP)
Filed on	23 August 1999 (23.08.1999)
US	60/159,849 (CIP)
Filed on	15 October 1999 (15.10.1999)

(71) Applicant *(for all designated States except US)*: **INCYTE GENOMICS, INC.** [US/US]; 3160 Porter Drive, Palo Alto, CA 94304 (US).

(72) Inventors; and
(75) Inventors/Applicants *(for US only)*: **YUE, Henry** [US/US]; 826 Lois Avenue, Sunnyvale, CA 94087 (US). **TANG, Y., Tom** [CN/US]; 4230 Ranwick Court, San Jose, CA 95118 (US). **BANDMAN, Olga** [US/US]; 366 Anna Avenue, Mountain View, CA 94043 (US). **HILLMAN, Jennifer, L.** [US/US]; 230 Monroe Drive #12, Montain View, CA 94040 (US). **LAL, Preeti** [IN/US]; 2382 Lass Drive, Santa Clara, CA 95054 (US). **AU-YOUNG, Janice** [US/US]; 233 Golden Eagle Lane, Brisbane, CA 94005 (US). **REDDY, Roopa** [IN/US]; 1233 W. McKinley Avenue, #3, Sunnyvale, CA 94086 (US). **YANG, Junming** [CN/US]; 7125 Bark Lane, San Jose, CA 95129 (US). **BAUGHN, Mariah, R.** [US/US]; 14244 Santiago Road, San Leandro, CA 94577 (US). **LU, Dyung, Aina, M.** [US/US]; 55 Park Belmont Place, San Jose, CA 95136 (US). **AZIMZAI, Yalda** [US/US]; 2045 Rock Springs Drive, Hayward, CA 94545 (US). **PATTERSON, Chandra** [US/US]; 490 Sherwood Way #1, Menlo Park, CA 94025 (US).

(74) Agents: **HAMLET-COX, Diana** et al.; Incyte Genomics, Inc., 3160 Porter Drive, Palo Alto, CA 94304 (US).

(81) Designated States *(national)*: AE, AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CR, CU, CZ, DE, DK, DM, EE, ES, FI, GB, GD, GE, GH, GM, HR, HU, ID, IL, IN, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MA, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, TZ, UA, UG, US, UZ, VN, YU, ZA, ZW.

(84) Designated States *(regional)*: ARIPO patent (GH, GM, KE, LS, MW, MZ, SD, SL, SZ, TZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published:
— *Without international search report and to be republished upon receipt of that report.*

For two-letter codes and other abbreviations, refer to the "Guidance Notes on Codes and Abbreviations" appearing at the beginning of each regular issue of the PCT Gazette.

WO 01/05970 A2

(54) Title: GTP-BINDING ASSOCIATED PROTEINS

(57) Abstract: The invention provides human GTP-binding associated proteins (GBAP) and polynucleotides which identify and encode GBAP. The invention also provides expression vectors, host cells, antibodies, agonists, and antagonists. The invention also provides methods for diagnosing, treating, or preventing disorders associated with expression of GBAP.

GTP-BINDING ASSOCIATED PROTEINS

TECHNICAL FIELD

This invention relates to nucleic acid and amino acid sequences of GTP-binding associated proteins and to the use of these sequences in the diagnosis, treatment, and prevention of immune system, reproductive, nervous system, and cell signaling disorders, and cell proliferative disorders including cancer.

BACKGROUND OF THE INVENTION

Guanine nucleotide binding proteins (GTP-binding proteins) are present in all eukaryotic cells and function in processes including metabolism, cellular growth, differentiation, signal transduction, cytoskeletal organization, and intracellular vesicle transport and secretion. In higher organisms they are involved in signaling that regulates such processes as the immune response (Aussel, C. et al. (1988) J. Immunol. 140:215-220), apoptosis, differentiation, and cell proliferation including oncogenesis (Dhanasekaran, N. et al. (1998) Oncogene 17:1383-1394).

The superfamily of GTP-binding proteins can be subdivided into groups such as translational factors, heterotrimeric GTP-binding proteins involved in transmembrane signaling processes (also called G-proteins), proto-oncogene Ras proteins, other low molecular weight GTP-binding proteins including the products of rab, rap, rho, rac, smg21, smg25, YPT, SEC4, and ARF genes, and tubulins (Kaziro, Y. et al. (1991) Annu. Rev. Biochem. 60:349-400).

GTP-binding proteins are involved in protein biosynthesis and include initiation factor 2 (IF-2), elongation factor 2 (EF-Tu), and elongation factor G (EF-G), observed in prokaryotes; and initiation factor 2 (eIF-2), elongation factor Iα (EF-Iα), elongation factor 2 (EF-2), and release factor 3 (eRF3) observed in eukaryotes (Kaziro, supra). IF-2 promotes the GTP-dependent binding of the tRNA to the small subunit of the ribosome, the step that initiates protein translation. Elongation factors promote the binding of tRNA and GTP and the displacement of GDP after hydrolysis as protein biosynthesis proceeds. eRF3 participates in the recognition of stop codons and the release of nascent proteins from ribosomes.

Heterotrimeric GTP-binding proteins are composed of 3 subunits (α, β and γ) which, in the resting state, associate as a trimer at the inner face of the plasma membrane. Heterotrimeric G-proteins may be classified based on the sequence similarity of α subunits into the Gs, Gi, Gq and G12 subgroups. In the resting state, the α subunit binds guanosine diphosphate (GDP), and stimulation of the G-protein by an activated receptor leads to exchange of GDP for guanosine triphosphate (GTP). This exchange results in the separation of the α from the β and γ subunits, which remain tightly

associated as a dimer. Both the α subunit and β-γ dimer are then able to interact with effectors, either individually or in a cooperative manner. The intrinsic GTPase activity of the α subunit hydrolyzes the bound GTP to GDP. This returns the α subunit to its inactive conformation and allows it to reassociate with the β-γ complex, thus restoring the system to its resting state (Kaziro, supra). Some α subunits show tissue-specific expression indicating a specialized signaling role (Dhanasekaran, supra).

The α-s class of G-protein subunits is sensitive to ADP-ribosylation by pertussis toxin which uncouples the receptor:G-protein interaction. This uncoupling blocks signal transduction to receptors that decrease cAMP levels. cAMP levels regulate ion channels and activate phospholipases. The inhibitory α-I class is also susceptible to modification by pertussis toxin, which prevents α-I from lowering cAMP levels. Two novel classes of α subunits refractory to pertussis toxin modification are α-q, which activates phospholipase C, and α-12, which has sequence homology with the Drosophila gene concertina and may contribute to the regulation of embryonic development (Simon, M.I. (1991) Science 252:802-808).

The mammalian G-protein β and γ subunits, each about 340 amino acids long, share more than 80% homology. The β subunit (also called β-transducin) contains seven repeating units, each about 43 amino acids long. This WD-repeat, or G-beta repeat motif, is found in a variety of proteins with regulatory function such as Sec13, a yeast WD repeat protein involved in vesicular traffic; coronin-2, a mammalian WD repeat protein involved in regulation of the actin cytoskeleton; and Bop1, a mammalian WD repeat protein involved in growth suppression (Garcia-Higuera, I. et al. (1998) J. Biol. Chem. 273:9041-9049; Okumura, M. et al. (1998) DNA Cell Biol. 17:779-787; Pestov, D.G. et al. (1998) Oncogene 17:3187-3197). The activity of the β and γ subunits may be regulated by other proteins such as calmodulin, phosducin, or the neural protein GAP 43 (Clapham, D.E. and E.J. Neer (1993) Nature 365:403-406). The β subunit sequences are highly conserved among species, suggesting that they perform a fundamentally important role in the organization and function of G-protein linked systems (Van der Voorn, L. and H.L. Ploegh (1992) FEBS Lett. 307:131-134).

Mutations and variant expression of β-transducin proteins are linked with various disorders. Mutations in LIS1, a subunit of the human platelet activating factor acetylhydrolase, cause Miller-Dieker lissencephaly. RACK1 binds activated protein kinase C, and RbAp48 binds retinoblastoma protein. CstF is required for polyadenylation of mammalian pre-mRNA in vitro and associates with subunits of cleavage-stimulating factor. Defects in the regulation of β-catenin contribute to the neoplastic transformation of human cells. The WD40 repeats of the human F-box protein βTrCP mediate binding to β-catenin, thus regulating the targeted degradation of β-catenin by ubiquitin ligase (Neer, E.J. et al. (1994) Nature 371:297-300; Hart, M. et al. (1999) Curr. Biol. 9:207-210).

The γ subunit sequences are more variable than those of the β subunits. They are often post-translationally modified by isoprenylation and carboxyl-methylation of a cysteine residue four amino

acids from the C-terminus. These modifications appear to be necessary for the interaction of the β-γ dimer with the membrane and with other GTP-binding proteins. The β-γ dimer has been shown to modulate the activity of adenylyl cyclase isoforms, phospholipase C, and some ion channels. It is involved in receptor phosphorylation via specific kinases and has been implicated in the p21ras-dependent activation of the MAP kinase cascade and the recognition of specific receptors by GTP-binding proteins (Clapham and Neer, supra).

G-proteins interact with a variety of effectors including adenylyl cyclase (Clapham and Neer, supra). The signaling pathway mediated by cAMP is mitogenic in hormone-dependent endocrine tissues such as adrenal cortex, thyroid, ovary, pituitary, and testes. Cancers in these tissues have been related to a mutationally activated form of a $G\alpha_s$ known as the gsp (Gs protein) oncogene (Dhanasekaran, supra). Another effector is phosducin, a retinal phosphoprotein, which forms a specific complex with retinal G-protein β and γ subunits and modulates the ability of the β-γ dimer to interact with retinal α subunits (Clapham and Neer, supra). Additional G-protein effectors include RIN1 (Ras interaction/interference), which acts as an effector of H-Ras and interferes with the Ras signal transduction pathway; Rabin3, which associates with the Ras-like GTPase Rab3A; and Rhotekin, a protein that binds with, and inhibits, Rho GTPase activity (Han, L. and J. Colicelli (1995) Mol. Cell Biol. 15:1318-1323; Brondyk, W.H. et al. (1995) Mol. Cell Biol. 15:1137-1143; and Reid, T. et al. (1996) J. Biol. Chem. 27:13556-13560).

The low molecular weight GTP-binding proteins regulate cell growth, cell cycle control, protein secretion, and intracellular vesicle interaction. These GTP-binding proteins respond to extracellular signals from receptors and activating proteins by transducing mitogenic signals (Tavitian, A. (1995) C. R. Seances Soc. Biol. Fil. 189:7-12). Low molecular weight GTP-binding proteins consist of single polypeptides of 21-30kD which, like the α subunit of heterotrimeric GTP-binding proteins, are able to bind to and hydrolyze GTP, thus cycling from an inactive to an active state. The intrinsic rate of GTP hydrolysis of these GTP-binding proteins is typically very slow, but it can be stimulated by several orders of magnitude by GTPase-activating proteins (GAPs), such as β2-chimaerin (Geyer, M. and Wittinghofer, A. (1997) Curr. Opin. Struct. Biol. 7:786-792; Caloca, M. J. et al. (1997) J. Biol. Chem. 272:26488-26496).

Low molecular weight GTP-binding proteins play critical roles in cellular protein trafficking events, such as the translocation of proteins and soluble complexes from the cytosol to the membrane through an exchange of GDP for GTP (Ktistakis, N.T. (1998) BioEssays 20:495-504). In vesicle transport, the interaction between vesicle- and target- specific identifiers (v-SNAREs and tSNAREs) docks the vesicle to the acceptor membrane. The budding process is regulated by GTPases such as the closely related ADP ribosylation factors (ARFs) and SAR proteins, while GTPases such as Rab allow assembly of SNARE complexes and may play a role in removal of defective complexes (Rothman, J.E.

and F.T. Wieland (1996) Science 272:227-234). The rab proteins control the translocation of vesicles to and from membranes for protein localization, protein processing, and secretion. The rho GTP-binding proteins control signal transduction pathways that link growth factor receptors to actin polymerization which is necessary for normal cellular growth and division. The ran GTP-binding proteins are located in the nucleus of cells and have a key role in nuclear protein import, the control of DNA synthesis, and cell-cycle progression (Hall, A. (1990) Science 249:635-640; Scheffzek, K. et al. (1995) Nature 374:378-381).

The Ras proteins Ras1, Ras2 and $G_s\alpha$ stimulate adenylyl cyclase (Kaziro, supra) which affects a broad array of cellular processes including determination of whether cells continue to grow or become terminally differentiated. Stimulation of cell surface receptors activates Ras which, in turn, activates cytoplasmic kinases. These kinases translocate to the nucleus and activate key transcription factors that control gene expression and protein synthesis (Barbacid, M. (1987) Annu. Rev. Biochem. 56:779-827; Treisman, R. (1994) Curr. Opin. Genet. Dev. 4:96-101). Mutant Ras-family proteins which bind but cannot hydrolyze GTP are permanently activated and are thus rendered oncogenic (Drivas, G.T. et al. (1990) Mol. Cell. Biol. 10:1793-1798).

Ras-like proteins have also been implicated in tumor suppression. For example, NOEY2, a novel gene encoding a Ras-like protein, is expressed in normal ovarian and breast epithelial cells. However, NOEY2 expression is reduced or abrogated in ovarian and breast carcinomas, suggesting a role for the NOEY2 gene product in tumor suppression (Yu, Y. et al. (1999) Proc. Natl. Acad. Sci. USA 96:214-219).

Irregularities in GTP-binding protein signaling cascades may result in abnormal activation of leukocytes and lymphocytes, leading to the tissue damage and destruction seen in many inflammatory and autoimmune diseases such as rheumatoid arthritis, biliary cirrhosis, hemolytic anemia, lupus erythematosus, and thyroiditis. Abnormal cell proliferation, including cyclic AMP-mediated stimulation of brain, thyroid, adrenal, and gonadal tissue proliferation is regulated by G proteins. Mutations in G_α subunits have been found in growth-hormone-secreting pituitary somatotroph tumors, hyperfunctioning thyroid adenomas, and ovarian and adrenal neoplasms (Meij, J.T.A. (1996) Mol. Cell. Biochem. 157:31-38; Aussel, supra).

The discovery of new GTP-binding associated proteins and the polynucleotides encoding them satisfies a need in the art by providing new compositions which are useful in the diagnosis, prevention, and treatment of immune system, reproductive, nervous system, and cell signaling disorders, and cell proliferative disorders including cancer.

SUMMARY OF THE INVENTION

The invention features purified polypeptides, GTP-binding associated proteins, referred to

collectively as "GBAP" and individually as "GBAP-1," "GBAP-2," "GBAP-3," "GBAP-4," "GBAP-5," "GBAP-6," "GBAP-7," "GBAP-8," "GBAP-9," "GBAP-10," "GBAP-11," "GBAP-12," "GBAP-13," "GBAP-14," "GBAP-15," "GBAP-16," "GBAP-17," "GBAP-18," "GBAP-19," "GBAP-20," "GBAP-21," "GBAP-22," "GBAP-23," "GBAP-24," "GBAP-25," "GBAP-26," "GBAP-27," "GBAP-28," "GBAP-29," "GBAP-30," "GBAP-31," "GBAP-32," "GBAP-33," "GBAP-34," "GBAP-35," "GBAP-36," "GBAP-37," "GBAP-38," "GBAP-39," "GBAP-40," "GBAP-41," "GBAP-42," "GBAP-43," "GBAP-44," "GBAP-45," "GBAP-46," "GBAP-47," "GBAP-48," "GBAP-49," "GBAP-50," "GBAP-51," "GBAP-52," "GBAP-53," "GBAP-54," "GBAP-55," "GBAP-56," "GBAP-57," "GBAP-58," "GBAP-59," "GBAP-60," "GBAP-61," "GBAP-62," "GBAP-63," "GBAP-64," "GBAP-65," and "GBAP-66." In one aspect, the invention provides an isolated polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66. In one alternative, the invention provides an isolated polypeptide comprising the amino acid sequence of SEQ ID NO:1-66.

The invention further provides an isolated polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66. In one alternative, the polynucleotide encodes a polypeptide selected from the group consisting of SEQ ID NO:1-66. In another alternative, the polynucleotide is selected from the group consisting of SEQ ID NO:67-132.

Additionally, the invention provides a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66. In one alternative, the invention provides a cell transformed with the recombinant polynucleotide. In another alternative, the invention provides a transgenic organism

comprising the recombinant polynucleotide.

The invention also provides a method for producing a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66. The method comprises a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide encoding the polypeptide, and b) recovering the polypeptide so expressed.

Additionally, the invention provides an isolated antibody which specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66.

The invention further provides an isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:67-132, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:67-132, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). In one alternative, the polynucleotide comprises at least 60 contiguous nucleotides.

Additionally, the invention provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:67-132, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:67-132, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or

fragments thereof, and b) detecting the presence or absence of said hybridization complex, and optionally, if present, the amount thereof. In one alternative, the probe comprises at least 60 contiguous nucleotides.

The invention further provides a method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:67-132, b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:67-132, c) a polynucleotide sequence complementary to a), d) a polynucleotide sequence complementary to b), and e) an RNA equivalent of a)-d). The method comprises a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

The invention further provides a composition comprising an effective amount of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, and a pharmaceutically acceptable excipient. In one embodiment, the l composition comprises an amino acid sequence selected from the group consisting of SEQ ID NO:1-66. The invention additionally provides a method of treating a disease or condition associated with decreased expression of functional GBAP, comprising administering to a patient in need of such treatment the composition.

The invention also provides a method for screening a compound for effectiveness as an agonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting agonist activity in the sample. In one alternative, the invention provides a composition comprising an agonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with decreased expression of

functional GBAP, comprising administering to a patient in need of such treatment the composition.

Additionally, the invention provides a method for screening a compound for effectiveness as an antagonist of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66. The method comprises a) exposing a sample comprising the polypeptide to a compound, and b) detecting antagonist activity in the sample. In one alternative, the invention provides a composition comprising an antagonist compound identified by the method and a pharmaceutically acceptable excipient. In another alternative, the invention provides a method of treating a disease or condition associated with overexpression of functional GBAP, comprising administering to a patient in need of such treatment the composition.

The invention further provides a method of screening for a compound that specifically binds to a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66. The method comprises a) combining the polypeptide with at least one test compound under suitable conditions, and b) detecting binding of the polypeptide to the test compound, thereby identifying a compound that specifically binds to the polypeptide.

The invention further provides a method of screening for a compound that modulates the activity of a polypeptide comprising an amino acid sequence selected from the group consisting of a) an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66, and d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1-66. The method comprises a) combining the polypeptide with at least one test compound under conditions permissive for the activity of the polypeptide, b) assessing the activity of the polypeptide in the presence of the test compound, and c) comparing the activity of the polypeptide in the presence of the test compound with the activity of the polypeptide in the absence of the test compound, wherein a

change in the activity of the polypeptide in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide.

The invention further provides a method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence selected from the group consisting of SEQ ID NO:67-132, the method comprising a) exposing a sample comprising the target polynucleotide to a compound, and b) detecting altered expression of the target polynucleotide.

The invention further provides a method for assessing toxicity of a test compound, said method comprising a) treating a biological sample containing nucleic acids with the test compound; b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide comprising a polynucleotide sequence selected from the group consisting of i) a polynucleotide sequence selected from the group consisting of SEQ ID NO:67-132, ii) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:67-132, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Hybridization occurs under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence selected from the group consisting of SEQ ID NO:67-132, ii) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:67-132, iii) a polynucleotide sequence complementary to i), iv) a polynucleotide sequence complementary to ii), and v) an RNA equivalent of i)-iv). Alternatively, the target polynucleotide comprises a fragment of the above polynucleotide sequence; c) quantifying the amount of hybridization complex; and d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

BRIEF DESCRIPTION OF THE TABLES

Table 1 shows polypeptide and nucleotide sequence identification numbers (SEQ ID NOs), clone identification numbers (clone IDs), cDNA libraries, and cDNA fragments used to assemble full-length sequences encoding GBAP.

Table 2 shows features of each polypeptide sequence, including potential motifs, homologous sequences, and methods, algorithms, and searchable databases used for analysis of GBAP.

Table 3 shows selected fragments of each nucleic acid sequence; the tissue-specific expression

patterns of each nucleic acid sequence as determined by northern analysis; diseases, disorders, or conditions associated with these tissues; and the vector into which each cDNA was cloned.

Table 4 describes the tissues used to construct the cDNA libraries from which cDNA clones encoding GBAP were isolated.

Table 5 shows the tools, programs, and algorithms used to analyze the polynucleotides and polypeptides of the invention, along with applicable descriptions, references, and threshold parameters.

DESCRIPTION OF THE INVENTION

Before the present proteins, nucleotide sequences, and methods are described, it is understood that this invention is not limited to the particular machines, materials and methods described, as these may vary. It is also to be understood that the terminology used herein is for the purpose of describing particular embodiments only, and is not intended to limit the scope of the present invention which will be limited only by the appended claims.

It must be noted that as used herein and in the appended claims, the singular forms "a," "an," and "the" include plural reference unless the context clearly dictates otherwise. Thus, for example, a reference to "a host cell" includes a plurality of such host cells, and a reference to "an antibody" is a reference to one or more antibodies and equivalents thereof known to those skilled in the art, and so forth.

Unless defined otherwise, all technical and scientific terms used herein have the same meanings as commonly understood by one of ordinary skill in the art to which this invention belongs. Although any machines, materials, and methods similar or equivalent to those described herein can be used to practice or test the present invention, the preferred machines, materials and methods are now described. All publications mentioned herein are cited for the purpose of describing and disclosing the cell lines, protocols, reagents and vectors which are reported in the publications and which might be used in connection with the invention. Nothing herein is to be construed as an admission that the invention is not entitled to antedate such disclosure by virtue of prior invention.

DEFINITIONS

"GBAP" refers to the amino acid sequences of substantially purified GBAP obtained from any species, particularly a mammalian species, including bovine, ovine, porcine, murine, equine, and human, and from any source, whether natural, synthetic, semi-synthetic, or recombinant.

The term "agonist" refers to a molecule which intensifies or mimics the biological activity of GBAP. Agonists may include proteins, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of GBAP either by directly interacting with GBAP or by acting on components of the biological pathway in which GBAP participates.

An "allelic variant" is an alternative form of the gene encoding GBAP. Allelic variants may

result from at least one mutation in the nucleic acid sequence and may result in altered mRNAs or in polypeptides whose structure or function may or may not be altered. A gene may have none, one, or many allelic variants of its naturally occurring form. Common mutational changes which give rise to allelic variants are generally ascribed to natural deletions, additions, or substitutions of nucleotides. Each of these types of changes may occur alone, or in combination with the others, one or more times in a given sequence.

"Altered" nucleic acid sequences encoding GBAP include those sequences with deletions, insertions, or substitutions of different nucleotides, resulting in a polypeptide the same as GBAP or a polypeptide with at least one functional characteristic of GBAP. Included within this definition are polymorphisms which may or may not be readily detectable using a particular oligonucleotide probe of the polynucleotide encoding GBAP, and improper or unexpected hybridization to allelic variants, with a locus other than the normal chromosomal locus for the polynucleotide sequence encoding GBAP. The encoded protein may also be "altered," and may contain deletions, insertions, or substitutions of amino acid residues which produce a silent change and result in a functionally equivalent GBAP. Deliberate amino acid substitutions may be made on the basis of similarity in polarity, charge, solubility, hydrophobicity, hydrophilicity, and/or the amphipathic nature of the residues, as long as the biological or immunological activity of GBAP is retained. For example, negatively charged amino acids may include aspartic acid and glutamic acid, and positively charged amino acids may include lysine and arginine. Amino acids with uncharged polar side chains having similar hydrophilicity values may include: asparagine and glutamine; and serine and threonine. Amino acids with uncharged side chains having similar hydrophilicity values may include: leucine, isoleucine, and valine; glycine and alanine; and phenylalanine and tyrosine.

The terms "amino acid" and "amino acid sequence" refer to an oligopeptide, peptide, polypeptide, or protein sequence, or a fragment of any of these, and to naturally occurring or synthetic molecules. Where "amino acid sequence" is recited to refer to a sequence of a naturally occurring protein molecule, "amino acid sequence" and like terms are not meant to limit the amino acid sequence to the complete native amino acid sequence associated with the recited protein molecule.

"Amplification" relates to the production of additional copies of a nucleic acid sequence. Amplification is generally carried out using polymerase chain reaction (PCR) technologies well known in the art.

The term "antagonist" refers to a molecule which inhibits or attenuates the biological activity of GBAP. Antagonists may include proteins such as antibodies, nucleic acids, carbohydrates, small molecules, or any other compound or composition which modulates the activity of GBAP either by directly interacting with GBAP or by acting on components of the biological pathway in which GBAP participates.

The term "antibody" refers to intact immunoglobulin molecules as well as to fragments thereof, such as Fab, $F(ab')_2$, and Fv fragments, which are capable of binding an epitopic determinant. Antibodies that bind GBAP polypeptides can be prepared using intact polypeptides or using fragments containing small peptides of interest as the immunizing antigen. The polypeptide or oligopeptide used to immunize an animal (e.g., a mouse, a rat, or a rabbit) can be derived from the translation of RNA, or synthesized chemically, and can be conjugated to a carrier protein if desired. Commonly used carriers that are chemically coupled to peptides include bovine serum albumin, thyroglobulin, and keyhole limpet hemocyanin (KLH). The coupled peptide is then used to immunize the animal.

The term "antigenic determinant" refers to that region of a molecule (i.e., an epitope) that makes contact with a particular antibody. When a protein or a fragment of a protein is used to immunize a host animal, numerous regions of the protein may induce the production of antibodies which bind specifically to antigenic determinants (particular regions or three-dimensional structures on the protein). An antigenic determinant may compete with the intact antigen (i.e., the immunogen used to elicit the immune response) for binding to an antibody.

The term "antisense" refers to any composition capable of base-pairing with the "sense" (coding) strand of a specific nucleic acid sequence. Antisense compositions may include DNA; RNA; peptide nucleic acid (PNA); oligonucleotides having modified backbone linkages such as phosphorothioates, methylphosphonates, or benzylphosphonates; oligonucleotides having modified sugar groups such as 2'-methoxyethyl sugars or 2'-methoxyethoxy sugars; or oligonucleotides having modified bases such as 5-methyl cytosine, 2'-deoxyuracil, or 7-deaza-2'-deoxyguanosine. Antisense molecules may be produced by any method including chemical synthesis or transcription. Once introduced into a cell, the complementary antisense molecule base-pairs with a naturally occurring nucleic acid sequence produced by the cell to form duplexes which block either transcription or translation. The designation "negative" or "minus" can refer to the antisense strand, and the designation "positive" or "plus" can refer to the sense strand of a reference DNA molecule.

The term "biologically active" refers to a protein having structural, regulatory, or biochemical functions of a naturally occurring molecule. Likewise, "immunologically active" or "immunogenic" refers to the capability of the natural, recombinant, or synthetic GBAP, or of any oligopeptide thereof, to induce a specific immune response in appropriate animals or cells and to bind with specific antibodies.

"Complementary" describes the relationship between two single-stranded nucleic acid sequences that anneal by base-pairing. For example, 5'-AGT-3' pairs with its complement, 3'-TCA-5'.

A "composition comprising a given polynucleotide sequence" and a "composition comprising a given amino acid sequence" refer broadly to any composition containing the given polynucleotide or

amino acid sequence. The composition may comprise a dry formulation or an aqueous solution. Compositions comprising polynucleotide sequences encoding GBAP or fragments of GBAP may be employed as hybridization probes. The probes may be stored in freeze-dried form and may be associated with a stabilizing agent such as a carbohydrate. In hybridizations, the probe may be deployed in an aqueous solution containing salts (e.g., NaCl), detergents (e.g., sodium dodecyl sulfate; SDS), and other components (e.g., Denhardt's solution, dry milk, salmon sperm DNA, etc.).

"Consensus sequence" refers to a nucleic acid sequence which has been subjected to repeated DNA sequence analysis to resolve uncalled bases, extended using the XL-PCR kit (PE Biosystems, Foster City CA) in the 5' and/or the 3' direction, and resequenced, or which has been assembled from one or more overlapping cDNA, EST, or genomic DNA fragments using a computer program for fragment assembly, such as the GELVIEW fragment assembly system (GCG, Madison WI) or Phrap (University of Washington, Seattle WA). Some sequences have been both extended and assembled to produce the consensus sequence.

"Conservative amino acid substitutions" are those substitutions that are predicted to least interfere with the properties of the original protein, i.e., the structure and especially the function of the protein is conserved and not significantly changed by such substitutions. The table below shows amino acids which may be substituted for an original amino acid in a protein and which are regarded as conservative amino acid substitutions.

Original Residue	Conservative Substitution
Ala	Gly, Ser
Arg	His, Lys
Asn	Asp, Gln, His
Asp	Asn, Glu
Cys	Ala, Ser
Gln	Asn, Glu, His
Glu	Asp, Gln, His
Gly	Ala
His	Asn, Arg, Gln, Glu
Ile	Leu, Val
Leu	Ile, Val
Lys	Arg, Gln, Glu
Met	Leu, Ile
Phe	His, Met, Leu, Trp, Tyr
Ser	Cys, Thr
Thr	Ser, Val
Trp	Phe, Tyr
Tyr	His, Phe, Trp
Val	Ile, Leu, Thr

Conservative amino acid substitutions generally maintain (a) the structure of the polypeptide backbone in the area of the substitution, for example, as a beta sheet or alpha helical conformation, (b) the charge or hydrophobicity of the molecule at the site of the substitution, and/or (c) the bulk of the

side chain.

. A "deletion" refers to a change in the amino acid or nucleotide sequence that results in the absence of one or more amino acid residues or nucleotides.

The term "derivative" refers to a chemically modified polynucleotide or polypeptide. Chemical modifications of a polynucleotide sequence can include, for example, replacement of hydrogen by an alkyl, acyl, hydroxyl, or amino group. A derivative polynucleotide encodes a polypeptide which retains at least one biological or immunological function of the natural molecule. A derivative polypeptide is one modified by glycosylation, pegylation, or any similar process that retains at least one biological or immunological function of the polypeptide from which it was derived.

A "detectable label" refers to a reporter molecule or enzyme that is capable of generating a measurable signal and is covalently or noncovalently joined to a polynucleotide or polypeptide.

A "fragment" is a unique portion of GBAP or the polynucleotide encoding GBAP which is identical in sequence to but shorter in length than the parent sequence. A fragment may comprise up to the entire length of the defined sequence, minus one nucleotide/amino acid residue. For example, a fragment may comprise from 5 to 1000 contiguous nucleotides or amino acid residues. A fragment used as a probe, primer, antigen, therapeutic molecule, or for other purposes, may be at least 5, 10, 15, 16, 20, 25, 30, 40, 50, 60, 75, 100, 150, 250 or at least 500 contiguous nucleotides or amino acid residues in length. Fragments may be preferentially selected from certain regions of a molecule. For example, a polypeptide fragment may comprise a certain length of contiguous amino acids selected from the first 250 or 500 amino acids (or first 25% or 50% of a polypeptide) as shown in a certain defined sequence. Clearly these lengths are exemplary, and any length that is supported by the specification, including the Sequence Listing, tables, and figures, may be encompassed by the present embodiments.

A fragment of SEQ ID NO:67-132 comprises a region of unique polynucleotide sequence that specifically identifies SEQ ID NO:67-132, for example, as distinct from any other sequence in the genome from which the fragment was obtained. A fragment of SEQ ID NO:67-132 is useful, for example, in hybridization and amplification technologies and in analogous methods that distinguish SEQ ID NO:67-132 from related polynucleotide sequences. The precise length of a fragment of SEQ ID NO:67-132 and the region of SEQ ID NO:67-132 to which the fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A fragment of SEQ ID NO:1-66 is encoded by a fragment of SEQ ID NO:67-132. A fragment of SEQ ID NO:1-66 comprises a region of unique amino acid sequence that specifically identifies SEQ ID NO:1-66. For example, a fragment of SEQ ID NO:1-66 is useful as an immunogenic peptide for the development of antibodies that specifically recognize SEQ ID NO:1-66. The precise length of a fragment of SEQ ID NO:1-66 and the region of SEQ ID NO:1-66 to which the

fragment corresponds are routinely determinable by one of ordinary skill in the art based on the intended purpose for the fragment.

A "full-length" polynucleotide sequence is one containing at least a translation initiation codon (e.g., methionine) followed by an open reading frame and a translation termination codon. A "full-length" polynucleotide sequence encodes a "full-length" polypeptide sequence.

"Homology" refers to sequence similarity or, interchangeably, sequence identity, between two or more polynucleotide sequences or two or more polypeptide sequences.

The terms "percent identity" and "% identity," as applied to polynucleotide sequences, refer to the percentage of residue matches between at least two polynucleotide sequences aligned using a standardized algorithm. Such an algorithm may insert, in a standardized and reproducible way, gaps in the sequences being compared in order to optimize alignment between two sequences, and therefore achieve a more meaningful comparison of the two sequences.

Percent identity between polynucleotide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program. This program is part of the LASERGENE software package, a suite of molecular biological analysis programs (DNASTAR, Madison WI). CLUSTAL V is described in Higgins, D.G. and P.M. Sharp (1989) CABIOS 5:151-153 and in Higgins, D.G. et al. (1992) CABIOS 8:189-191. For pairwise alignments of polynucleotide sequences, the default parameters are set as follows: Ktuple=2, gap penalty=5, window=4, and "diagonals saved"=4. The "weighted" residue weight table is selected as the default. Percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polynucleotide sequences.

Alternatively, a suite of commonly used and freely available sequence comparison algorithms is provided by the National Center for Biotechnology Information (NCBI) Basic Local Alignment Search Tool (BLAST) (Altschul, S.F. et al. (1990) J. Mol. Biol. 215:403-410), which is available from several sources, including the NCBI, Bethesda, MD, and on the Internet at http://www.ncbi.nlm.nih.gov/BLAST/. The BLAST software suite includes various sequence analysis programs including "blastn," that is used to align a known polynucleotide sequence with other polynucleotide sequences from a variety of databases. Also available is a tool called "BLAST 2 Sequences" that is used for direct pairwise comparison of two nucleotide sequences. "BLAST 2 Sequences" can be accessed and used interactively at http://www.ncbi.nlm.nih.gov/gorf/bl2.html. The "BLAST 2 Sequences" tool can be used for both blastn and blastp (discussed below). BLAST programs are commonly used with gap and other parameters set to default settings. For example, to compare two nucleotide sequences, one may use blastn with the "BLAST 2 Sequences" tool Version 2.0.12 (April-21-2000) set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Reward for match: 1

Penalty for mismatch: -2

Open Gap: 5 and Extension Gap: 2 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 11

Filter: on

Percent identity may be measured over the length of an entire defined sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined sequence, for instance, a fragment of at least 20, at least 30, at least 40, at least 50, at least 70, at least 100, or at least 200 contiguous nucleotides. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures, or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

Nucleic acid sequences that do not show a high degree of identity may nevertheless encode similar amino acid sequences due to the degeneracy of the genetic code. It is understood that changes in a nucleic acid sequence can be made using this degeneracy to produce multiple nucleic acid sequences that all encode substantially the same protein.

The phrases "percent identity" and "% identity," as applied to polypeptide sequences, refer to the percentage of residue matches between at least two polypeptide sequences aligned using a standardized algorithm. Methods of polypeptide sequence alignment are well-known. Some alignment methods take into account conservative amino acid substitutions. Such conservative substitutions, explained in more detail above, generally preserve the charge and hydrophobicity at the site of substitution, thus preserving the structure (and therefore function) of the polypeptide.

Percent identity between polypeptide sequences may be determined using the default parameters of the CLUSTAL V algorithm as incorporated into the MEGALIGN version 3.12e sequence alignment program (described and referenced above). For pairwise alignments of polypeptide sequences using CLUSTAL V, the default parameters are set as follows: Ktuple=1, gap penalty=3, window=5, and "diagonals saved"=5. The PAM250 matrix is selected as the default residue weight table. As with polynucleotide alignments, the percent identity is reported by CLUSTAL V as the "percent similarity" between aligned polypeptide sequence pairs.

Alternatively the NCBI BLAST software suite may be used. For example, for a pairwise comparison of two polypeptide sequences, one may use the "BLAST 2 Sequences" tool Version 2.0.12 (Apr-21-2000) with blastp set at default parameters. Such default parameters may be, for example:

Matrix: BLOSUM62

Open Gap: 11 and Extension Gap: 1 penalties

Gap x drop-off: 50

Expect: 10

Word Size: 3

Filter: on

Percent identity may be measured over the length of an entire defined polypeptide sequence, for example, as defined by a particular SEQ ID number, or may be measured over a shorter length, for example, over the length of a fragment taken from a larger, defined polypeptide sequence, for instance, a fragment of at least 15, at least 20, at least 30, at least 40, at least 50, at least 70 or at least 150 contiguous residues. Such lengths are exemplary only, and it is understood that any fragment length supported by the sequences shown herein, in the tables, figures or Sequence Listing, may be used to describe a length over which percentage identity may be measured.

"Human artificial chromosomes" (HACs) are linear microchromosomes which may contain DNA sequences of about 6 kb to 10 Mb in size, and which contain all of the elements required for chromosome replication, segregation and maintenance.

The term "humanized antibody" refers to an antibody molecule in which the amino acid sequence in the non-antigen binding regions has been altered so that the antibody more closely resembles a human antibody, and still retains its original binding ability.

"Hybridization" refers to the process by which a polynucleotide strand anneals with a complementary strand through base pairing under defined hybridization conditions. Specific hybridization is an indication that two nucleic acid sequences share a high degree of complementarity. Specific hybridization complexes form under permissive annealing conditions and remain hybridized after the "washing" step(s). The washing step(s) is particularly important in determining the stringency of the hybridization process, with more stringent conditions allowing less non-specific binding, i.e., binding between pairs of nucleic acid strands that are not perfectly matched. Permissive conditions for annealing of nucleic acid sequences are routinely determinable by one of ordinary skill in the art and may be consistent among hybridization experiments, whereas wash conditions may be varied among experiments to achieve the desired stringency, and therefore hybridization specificity. Permissive annealing conditions occur, for example, at 68°C in the presence of about 6 x SSC, about 1% (w/v) SDS, and about 100 μg/ml sheared, denatured salmon sperm DNA.

Generally, stringency of hybridization is expressed, in part, with reference to the temperature under which the wash step is carried out. Such wash temperatures are typically selected to be about 5°C to 20°C lower than the thermal melting point (T_m) for the specific sequence at a defined ionic strength and pH. The T_m is the temperature (under defined ionic strength and pH) at which 50% of the target sequence hybridizes to a perfectly matched probe. An equation for calculating T_m and conditions

for nucleic acid hybridization are well known and can be found in Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; specifically see volume 2, chapter 9.

High stringency conditions for hybridization between polynucleotides of the present invention include wash conditions of 68°C in the presence of about 0.2 x SSC and about 0.1% SDS, for 1 hour. Alternatively, temperatures of about 65°C, 60°C, 55°C, or 42°C may be used. SSC concentration may be varied from about 0.1 to 2 x SSC, with SDS being present at about 0.1%. Typically, blocking reagents are used to block non-specific hybridization. Such blocking reagents include, for instance, sheared and denatured salmon sperm DNA at about 100-200 μg/ml. Organic solvent, such as formamide at a concentration of about 35-50% v/v, may also be used under particular circumstances, such as for RNA:DNA hybridizations. Useful variations on these wash conditions will be readily apparent to those of ordinary skill in the art. Hybridization, particularly under high stringency conditions, may be suggestive of evolutionary similarity between the nucleotides. Such similarity is strongly indicative of a similar role for the nucleotides and their encoded polypeptides.

The term "hybridization complex" refers to a complex formed between two nucleic acid sequences by virtue of the formation of hydrogen bonds between complementary bases. A hybridization complex may be formed in solution (e.g., C_0t or R_0t analysis) or formed between one nucleic acid sequence present in solution and another nucleic acid sequence immobilized on a solid support (e.g., paper, membranes, filters, chips, pins or glass slides, or any other appropriate substrate to which cells or their nucleic acids have been fixed).

The words "insertion" and "addition" refer to changes in an amino acid or nucleotide sequence resulting in the addition of one or more amino acid residues or nucleotides, respectively.

"Immune response" can refer to conditions associated with inflammation, trauma, immune disorders, or infectious or genetic disease, etc. These conditions can be characterized by expression of various factors, e.g., cytokines, chemokines, and other signaling molecules, which may affect cellular and systemic defense systems.

An "immunogenic fragment" is a polypeptide or oligopeptide fragment of GBAP which is capable of eliciting an immune response when introduced into a living organism, for example, a mammal. The term "immunogenic fragment" also includes any polypeptide or oligopeptide fragment of GBAP which is useful in any of the antibody production methods disclosed herein or known in the art.

The term "microarray" refers to an arrangement of a plurality of polynucleotides, polypeptides, or other chemical compounds on a substrate.

The terms "element" and "array element" refer to a polynucleotide, polypeptide, or other chemical compound having a unique and defined position on a microarray.

The term "modulate" refers to a change in the activity of GBAP. For example, modulation

may cause an increase or a decrease in protein activity, binding characteristics, or any other biological, functional, or immunological properties of GBAP.

The phrases "nucleic acid" and "nucleic acid sequence" refer to a nucleotide, oligonucleotide, polynucleotide, or any fragment thereof. These phrases also refer to DNA or RNA of genomic or synthetic origin which may be single-stranded or double-stranded and may represent the sense or the antisense strand, to peptide nucleic acid (PNA), or to any DNA-like or RNA-like material.

"Operably linked" refers to the situation in which a first nucleic acid sequence is placed in a functional relationship with a second nucleic acid sequence. For instance, a promoter is operably linked to a coding sequence if the promoter affects the transcription or expression of the coding sequence. Operably linked DNA sequences may be in close proximity or contiguous and, where necessary to join two protein coding regions, in the same reading frame.

"Peptide nucleic acid" (PNA) refers to an antisense molecule or anti-gene agent which comprises an oligonucleotide of at least about 5 nucleotides in length linked to a peptide backbone of amino acid residues ending in lysine. The terminal lysine confers solubility to the composition. PNAs preferentially bind complementary single stranded DNA or RNA and stop transcript elongation, and may be pegylated to extend their lifespan in the cell.

"Post-translational modification" of an GBAP may involve lipidation, glycosylation, phosphorylation, acetylation, racemization, proteolytic cleavage, and other modifications known in the art. These processes may occur synthetically or biochemically. Biochemical modifications will vary by cell type depending on the enzymatic milieu of GBAP.

"Probe" refers to nucleic acid sequences encoding GBAP, their complements, or fragments thereof, which are used to detect identical, allelic or related nucleic acid sequences. Probes are isolated oligonucleotides or polynucleotides attached to a detectable label or reporter molecule. Typical labels include radioactive isotopes, ligands, chemiluminescent agents, and enzymes. "Primers" are short nucleic acids, usually DNA oligonucleotides, which may be annealed to a target polynucleotide by complementary base-pairing. The primer may then be extended along the target DNA strand by a DNA polymerase enzyme. Primer pairs can be used for amplification (and identification) of a nucleic acid sequence, e.g., by the polymerase chain reaction (PCR).

Probes and primers as used in the present invention typically comprise at least 15 contiguous nucleotides of a known sequence. In order to enhance specificity, longer probes and primers may also be employed, such as probes and primers that comprise at least 20, 25, 30, 40, 50, 60, 70, 80, 90, 100, or at least 150 consecutive nucleotides of the disclosed nucleic acid sequences. Probes and primers may be considerably longer than these examples, and it is understood that any length supported by the specification, including the tables, figures, and Sequence Listing, may be used.

Methods for preparing and using probes and primers are described in the references, for

example Sambrook, J. et al., 1989, Molecular Cloning: A Laboratory Manual, 2nd ed., vol. 1-3, Cold Spring Harbor Press, Plainview NY; Ausubel, F.M. et al.,1987, Current Protocols in Molecular Biology, Greene Publ. Assoc. & Wiley-Intersciences, New York NY; Innis, M. et al., 1990, PCR Protocols, A Guide to Methods and Applications, Academic Press, San Diego CA. PCR primer pairs can be derived from a known sequence, for example, by using computer programs intended for that purpose such as Primer (Version 0.5, 1991, Whitehead Institute for Biomedical Research, Cambridge MA).

Oligonucleotides for use as primers are selected using software known in the art for such purpose. For example, OLIGO 4.06 software is useful for the selection of PCR primer pairs of up to 100 nucleotides each, and for the analysis of oligonucleotides and larger polynucleotides of up to 5,000 nucleotides from an input polynucleotide sequence of up to 32 kilobases. Similar primer selection programs have incorporated additional features for expanded capabilities. For example, the PrimOU primer selection program (available to the public from the Genome Center at University of Texas South West Medical Center, Dallas TX) is capable of choosing specific primers from megabase sequences and is thus useful for designing primers on a genome-wide scope. The Primer3 primer selection program (available to the public from the Whitehead Institute/MIT Center for Genome Research, Cambridge MA) allows the user to input a "mispriming library," in which sequences to avoid as primer binding sites are user-specified. Primer3 is useful, in particular, for the selection of oligonucleotides for microarrays. (The source code for the latter two primer selection programs may also be obtained from their respective sources and modified to meet the user's specific needs.) The PrimeGen program (available to the public from the UK Human Genome Mapping Project Resource Centre, Cambridge UK) designs primers based on multiple sequence alignments, thereby allowing selection of primers that hybridize to either the most conserved or least conserved regions of aligned nucleic acid sequences. Hence, this program is useful for identification of both unique and conserved oligonucleotides and polynucleotide fragments. The oligonucleotides and polynucleotide fragments identified by any of the above selection methods are useful in hybridization technologies, for example, as PCR or sequencing primers, microarray elements, or specific probes to identify fully or partially complementary polynucleotides in a sample of nucleic acids. Methods of oligonucleotide selection are not limited to those described above.

A "recombinant nucleic acid" is a sequence that is not naturally occurring or has a sequence that is made by an artificial combination of two or more otherwise separated segments of sequence. This artificial combination is often accomplished by chemical synthesis or, more commonly, by the artificial manipulation of isolated segments of nucleic acids, e.g., by genetic engineering techniques such as those described in Sambrook, supra. The term recombinant includes nucleic acids that have been altered solely by addition, substitution, or deletion of a portion of the nucleic acid. Frequently, a

recombinant nucleic acid may include a nucleic acid sequence operably linked to a promoter sequence. Such a recombinant nucleic acid may be part of a vector that is used, for example, to transform a cell.

Alternatively, such recombinant nucleic acids may be part of a viral vector, e.g., based on a vaccinia virus, that could be use to vaccinate a mammal wherein the recombinant nucleic acid is expressed, inducing a protective immunological response in the mammal.

A "regulatory element" refers to a nucleic acid sequence usually derived from untranslated regions of a gene and includes enhancers, promoters, introns, and 5' and 3' untranslated regions (UTRs). Regulatory elements interact with host or viral proteins which control transcription, translation, or RNA stability.

"Reporter molecules" are chemical or biochemical moieties used for labeling a nucleic acid, amino acid, or antibody. Reporter molecules include radionuclides; enzymes; fluorescent, chemiluminescent, or chromogenic agents; substrates; cofactors; inhibitors; magnetic particles; and other moieties known in the art.

An "RNA equivalent," in reference to a DNA sequence, is composed of the same linear sequence of nucleotides as the reference DNA sequence with the exception that all occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The term "sample" is used in its broadest sense. A sample suspected of containing nucleic acids encoding GBAP, or fragments thereof, or GBAP itself, may comprise a bodily fluid; an extract from a cell, chromosome, organelle, or membrane isolated from a cell; a cell; genomic DNA, RNA, or cDNA, in solution or bound to a substrate; a tissue; a tissue print; etc.

The terms "specific binding" and "specifically binding" refer to that interaction between a protein or peptide and an agonist, an antibody, an antagonist, a small molecule, or any natural or synthetic binding composition. The interaction is dependent upon the presence of a particular structure of the protein, e.g., the antigenic determinant or epitope, recognized by the binding molecule. For example, if an antibody is specific for epitope "A," the presence of a polypeptide comprising the epitope A, or the presence of free unlabeled A, in a reaction containing free labeled A and the antibody will reduce the amount of labeled A that binds to the antibody.

The term "substantially purified" refers to nucleic acid or amino acid sequences that are removed from their natural environment and are isolated or separated, and are at least 60% free, preferably at least 75% free, and most preferably at least 90% free from other components with which they are naturally associated.

A "substitution" refers to the replacement of one or more amino acid residues or nucleotides by different amino acid residues or nucleotides, respectively.

"Substrate" refers to any suitable rigid or semi-rigid support including membranes, filters,

chips, slides, wafers, fibers, magnetic or nonmagnetic beads, gels, tubing, plates, polymers, microparticles and capillaries. The substrate can have a variety of surface forms, such as wells, trenches, pins, channels and pores, to which polynucleotides or polypeptides are bound.

A "transcript image" refers to the collective pattern of gene expression by a particular cell type or tissue under given conditions at a given time.

"Transformation" describes a process by which exogenous DNA is introduced into a recipient cell. Transformation may occur under natural or artificial conditions according to various methods well known in the art, and may rely on any known method for the insertion of foreign nucleic acid sequences into a prokaryotic or eukaryotic host cell. The method for transformation is selected based on the type of host cell being transformed and may include, but is not limited to, bacteriophage or viral infection, electroporation, heat shock, lipofection, and particle bombardment. The term "transformed" cells includes stably transformed cells in which the inserted DNA is capable of replication either as an autonomously replicating plasmid or as part of the host chromosome, as well as transiently transformed cells which express the inserted DNA or RNA for limited periods of time.

A "transgenic organism," as used herein, is any organism, including but not limited to animals and plants, in which one or more of the cells of the organism contains heterologous nucleic acid introduced by way of human intervention, such as by transgenic techniques well known in the art. The nucleic acid is introduced into the cell, directly or indirectly by introduction into a precursor of the cell, by way of deliberate genetic manipulation, such as by microinjection or by infection with a recombinant virus. The term genetic manipulation does not include classical cross-breeding, or in vitro fertilization, but rather is directed to the introduction of a recombinant DNA molecule. The transgenic organisms contemplated in accordance with the present invention include bacteria, cyanobacteria, fungi, plants, and animals. The isolated DNA of the present invention can be introduced into the host by methods known in the art, for example infection, transfection, transformation or transconjugation. Techniques for transferring the DNA of the present invention into such organisms are widely known and provided in references such as Sambrook et al. (1989), supra.

A "variant" of a particular nucleic acid sequence is defined as a nucleic acid sequence having at least 40% sequence identity to the particular nucleic acid sequence over a certain length of one of the nucleic acid sequences using blastn with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of nucleic acids may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 85%, at least 90%, at least 95% or at least 98% or greater sequence identity over a certain defined length. A variant may be described as, for example, an "allelic" (as defined above), "splice," "species," or "polymorphic" variant. A splice variant may have significant identity to a reference molecule, but will generally have a greater or lesser number of polynucleotides

due to alternative splicing of exons during mRNA processing. The corresponding polypeptide may possess additional functional domains or lack domains that are present in the reference molecule. Species variants are polynucleotide sequences that vary from one species to another. The resulting polypeptides generally will have significant amino acid identity relative to each other. A polymorphic variant is a variation in the polynucleotide sequence of a particular gene between individuals of a given species. Polymorphic variants also may encompass "single nucleotide polymorphisms" (SNPs) in which the polynucleotide sequence varies by one nucleotide base. The presence of SNPs may be indicative of, for example, a certain population, a disease state, or a propensity for a disease state.

A "variant" of a particular polypeptide sequence is defined as a polypeptide sequence having at least 40% sequence identity to the particular polypeptide sequence over a certain length of one of the polypeptide sequences using blastp with the "BLAST 2 Sequences" tool Version 2.0.9 (May-07-1999) set at default parameters. Such a pair of polypeptides may show, for example, at least 50%, at least 60%, at least 70%, at least 80%, at least 90%, at least 95%, or at least 98% or greater sequence identity over a certain defined length of one of the polypeptides.

THE INVENTION

The invention is based on the discovery of new human GTP-binding associated proteins (GBAP), the polynucleotides encoding GBAP, and the use of these compositions for the diagnosis, treatment, or prevention of immune system, reproductive, nervous system, and cell signaling disorders, and cell proliferative disorders including cancer.

Table 1 lists the Incyte clones used to assemble full length nucleotide sequences encoding GBAP. Columns 1 and 2 show the sequence identification numbers (SEQ ID NOs) of the polypeptide and nucleotide sequences, respectively. Column 3 shows the clone IDs of the Incyte clones in which nucleic acids encoding each GBAP were identified, and column 4 shows the cDNA libraries from which these clones were isolated. Column 5 shows Incyte clones and their corresponding cDNA libraries. Clones for which cDNA libraries are not indicated were derived from pooled cDNA libraries. In some cases, GenBank sequence identifiers are also shown in column 5. The Incyte clones and GenBank cDNA sequences, where indicated, in column 5 were used to assemble the consensus nucleotide sequence of each GBAP and are useful as fragments in hybridization technologies.

The columns of Table 2 show various properties of each of the polypeptides of the invention: column 1 references the SEQ ID NO; column 2 shows the number of amino acid residues in each polypeptide; column 3 shows potential phosphorylation sites; column 4 shows potential glycosylation sites; column 5 shows the amino acid residues comprising signature sequences and motifs; column 6 shows homologous sequences as identified by BLAST analysis; and column 7 shows analytical methods and in some cases, searchable databases to which the analytical methods were applied. The methods of column 7 were used to characterize each polypeptide through sequence homology and protein motifs.

The columns of Table 3 show the tissue-specificity and diseases, disorders, or conditions associated with nucleotide sequences encoding GBAP. The first column of Table 3 lists the nucleotide SEQ ID NOs. Column 2 lists fragments of the nucleotide sequences of column 1. These fragments are useful, for example, in hybridization or amplification technologies to identify SEQ ID NO:67-132 and to distinguish between SEQ ID NO:67-132 and related polynucleotide sequences. The polypeptides encoded by these fragments are useful, for example, as immunogenic peptides. Column 3 lists tissue categories which express GBAP as a fraction of total tissues expressing GBAP. Column 4 lists diseases, disorders, or conditions associated with those tissues expressing GBAP as a fraction of total tissues expressing GBAP. Column 5 lists the vectors used to subclone each cDNA library. Of particular note is the expression of SEQ ID NO:84 in lung tissues, and the tissue-specific expression of SEQ ID NO:132. Over 90% of tissues expressing SEQ ID NO:132 are derived from the nervous system, particularly the brain.

The columns of Table 4 show descriptions of the tissues used to construct the cDNA libraries from which cDNA clones encoding GBAP were isolated. Column 1 references the nucleotide SEQ ID NOs, column 2 shows the cDNA libraries from which these clones were isolated, and column 3 shows the tissue origins and other descriptive information relevant to the cDNA libraries in column 2.

SEQ ID NO:70 maps to chromosome 7 within the interval from 111.6 to 123.4 centiMorgans. This interval contains a gene that is down regulated in adenoma. SEQ ID NO:74 maps to chromosome 11 within the interval from 104.8 to 123.5 centiMorgans. This interval contains a gene associated with the cerebellar degenerative disorder, ataxia telangiectasia. SEQ ID NO:75 maps to chromosome 17 within the interval from 62.9 to 65.0 centiMorgans. SEQ ID NO:77 maps to chromosome 3 within the interval from 12.9 to 16.5 centiMorgans. SEQ ID NO:80 maps to chromosome 9 within the interval from 42.0 to 57.3 centiMorgans. SEQ ID NO:86 maps to chromosome 1 within the interval from 159.6 to 164.1 centiMorgans. SEQ ID NO:87 maps to chromosome 11 within the interval from 147.2 to 151.6. SEQ ID NO:90 maps to chromosome 1 within the interval from 219.2 to 223.0 centiMorgans. This interval contains a gene encoding a RAB interacting protein. SEQ ID NO:92 and SEQ ID NO:106 both map to chromosome 1 within the interval from 48.8 to 81.6 centiMorgans. This interval also contains genes associated with familial hypercholesterolemia, glucose transport defect, infantile hypophosphatasia, infantile neuronal ceroid lipofuscinosis, Kostmann disease, multiple epiphyseal dysplasia, porphyria cutanea tarda, and T-cell acute lymphocytic leukemia 1. SEQ ID NO:93 maps to chromosome 12 within the interval from 76.5 to 87.6 centiMorgans. This interval also contains genes associated with mucopolysaccharidosis type IIID, pseudovitamin D deficiency rickets, and renal amyloidosis. SEQ ID NO:94 and SEQ ID NO:109 both map to chromosome 1 within the interval from 143.1 to 146.6 centiMorgans, to chromosome 14 within the interval from 46.8 to 50.9 centiMorgans, to chromosome 16 within the interval from 88.1 to 90.2 centiMorgans, and to chromosome 19 within the

interval from 58.7 to 97.5 centiMorgans. The interval on chromosome 14 from 46.8 to 50.9 centiMorgans also contains a gene associated with dopa-responsive dystonia. The interval on chromosome 19 from 58.7 to 97.5 centiMorgans also contains genes associated with colorectal cancer, DNA ligase I deficiency, glutaricaciduria IIB, myotonic dystrophy, renal amyloidosis, T-cell acute lymphoblastic leukemia, and xeroderma pigmentosum D. SEQ ID NO:97 maps to chromosome 2 within the interval from 236.2 to 269.5 centiMorgans. This interval also contains genes associated with Crigler-Najjar syndrome, familial hypercholesterolemia, Oguchi disease, and primary hyperoxaluria. SEQ ID NO:101 maps to chromosome 2 within the interval from 225.6 to 233.1 centiMorgans, to chromosome 6 within the interval from 132.7 to 144.4 centiMorgans, and to chromosome 11 within the interval from 117.9 to 120.8 centiMorgans. The interval on chromosome 2 from 225.6 to 233.1 centiMorgans also contains a gene associated with Waardenburg syndrome 1. The interval on chromosome 6 from 132.7 to 144.4 centiMorgans also contains genes associated with familial disseminated atypical mycobacterial infection and rhizomelic chondrodysplasia punctata. The interval on chromosome 11 from 117.9 to 120.8 centiMorgans also contains a gene associated with acute intermittent porphyria. SEQ ID NO:111 maps to chromosome 19 within the interval from 35.5 to 49.4 centiMorgans, to chromosome 1 within the interval from the p-terminus to 16.4 centiMorgans, and to chromosome 11 within the interval from 147.2 centiMorgans to the q-terminus. SEQ ID NO:112 maps to chromosome 19 within the interval from 41.7 to 49.4 centiMorgans. SEQ ID NO:113 maps to chromosome 9 within the interval from 136.2 to 163.0 centiMorgans. SEQ ID NO:115 maps to chromosome 14 within the interval from 95.5 to 103.7 centiMorgans and to the X chromosome (23) within the interval from the p-terminus to 55.5 centiMorgans. SEQ ID NO:117 maps to chromosome 13 at 46.9 centiMorgans. SEQ ID NO:118 maps to chromosome 1 within the interval from 16.4 to 22.9 centiMorgans. SEQ ID NO:121 maps to chromosome 12 within the interval from 116.6 to 118.9 centiMorgans. SEQ ID NO:128 maps to chromosome 1 within the interval from the p-terminus to 16.4 centiMorgans.

The invention also encompasses GBAP variants. A preferred GBAP variant is one which has at least about 80%, or alternatively at least about 90%, or even at least about 95% amino acid sequence identity to the GBAP amino acid sequence, and which contains at least one functional or structural characteristic of GBAP.

The invention also encompasses polynucleotides which encode GBAP. In a particular embodiment, the invention encompasses a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:67-132, which encodes GBAP. The polynucleotide sequences of SEQ ID NO:67-132, as presented in the Sequence Listing, embrace the equivalent RNA sequences, wherein occurrences of the nitrogenous base thymine are replaced with uracil, and the sugar backbone is composed of ribose instead of deoxyribose.

The invention also encompasses a variant of a polynucleotide sequence encoding GBAP. In particular, such a variant polynucleotide sequence will have at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to the polynucleotide sequence encoding GBAP. A particular aspect of the invention encompasses a variant of a polynucleotide sequence comprising a sequence selected from the group consisting of SEQ ID NO:67-132 which has at least about 70%, or alternatively at least about 85%, or even at least about 95% polynucleotide sequence identity to a nucleic acid sequence selected from the group consisting of SEQ ID NO:67-132. Any one of the polynucleotide variants described above can encode an amino acid sequence which contains at least one functional or structural characteristic of GBAP.

It will be appreciated by those skilled in the art that as a result of the degeneracy of the genetic code, a multitude of polynucleotide sequences encoding GBAP, some bearing minimal similarity to the polynucleotide sequences of any known and naturally occurring gene, may be produced. Thus, the invention contemplates each and every possible variation of polynucleotide sequence that could be made by selecting combinations based on possible codon choices. These combinations are made in accordance with the standard triplet genetic code as applied to the polynucleotide sequence of naturally occurring GBAP, and all such variations are to be considered as being specifically disclosed.

Although nucleotide sequences which encode GBAP and its variants are generally capable of hybridizing to the nucleotide sequence of the naturally occurring GBAP under appropriately selected conditions of stringency, it may be advantageous to produce nucleotide sequences encoding GBAP or its derivatives possessing a substantially different codon usage, e.g., inclusion of non-naturally occurring codons. Codons may be selected to increase the rate at which expression of the peptide occurs in a particular prokaryotic or eukaryotic host in accordance with the frequency with which particular codons are utilized by the host. Other reasons for substantially altering the nucleotide sequence encoding GBAP and its derivatives without altering the encoded amino acid sequences include the production of RNA transcripts having more desirable properties, such as a greater half-life, than transcripts produced from the naturally occurring sequence.

The invention also encompasses production of DNA sequences which encode GBAP and GBAP derivatives, or fragments thereof, entirely by synthetic chemistry. After production, the synthetic sequence may be inserted into any of the many available expression vectors and cell systems using reagents well known in the art. Moreover, synthetic chemistry may be used to introduce mutations into a sequence encoding GBAP or any fragment thereof.

Also encompassed by the invention are polynucleotide sequences that are capable of hybridizing to the claimed polynucleotide sequences, and, in particular, to those shown in SEQ ID NO:67-132 and fragments thereof under various conditions of stringency. (See, e.g., Wahl, G.M. and S.L. Berger (1987) Methods Enzymol. 152:399-407; Kimmel, A.R. (1987) Methods Enzymol.

152:507-511.) Hybridization conditions, including annealing and wash conditions, are described in "Definitions."

Methods for DNA sequencing are well known in the art and may be used to practice any of the embodiments of the invention. The methods may employ such enzymes as the Klenow fragment of DNA polymerase I, SEQUENASE (US Biochemical, Cleveland OH), Taq polymerase (PE Biosystems, Foster City CA), thermostable T7 polymerase (Amersham Pharmacia Biotech, Piscataway NJ), or combinations of polymerases and proofreading exonucleases such as those found in the ELONGASE amplification system (Life Technologies, Gaithersburg MD). Preferably, sequence preparation is automated with machines such as the MICROLAB 2200 liquid transfer system (Hamilton, Reno NV), PTC200 thermal cycler (MJ Research, Watertown MA) and ABI CATALYST 800 thermal cycler (PE Biosystems). Sequencing is then carried out using either the ABI 373 or 377 DNA sequencing system (PE Biosystems), the MEGABACE 1000 DNA sequencing system (Molecular Dynamics, Sunnyvale CA), or other systems known in the art. The resulting sequences are analyzed using a variety of algorithms which are well known in the art. (See, e.g., Ausubel, F.M. (1997) Short Protocols in Molecular Biology, John Wiley & Sons, New York NY, unit 7.7; Meyers, R.A. (1995) Molecular Biology and Biotechnology, Wiley VCH, New York NY, pp. 856-853.)

The nucleic acid sequences encoding GBAP may be extended utilizing a partial nucleotide sequence and employing various PCR-based methods known in the art to detect upstream sequences, such as promoters and regulatory elements. For example, one method which may be employed, restriction-site PCR, uses universal and nested primers to amplify unknown sequence from genomic DNA within a cloning vector. (See, e.g., Sarkar, G. (1993) PCR Methods Applic. 2:318-322.) Another method, inverse PCR, uses primers that extend in divergent directions to amplify unknown sequence from a circularized template. The template is derived from restriction fragments comprising a known genomic locus and surrounding sequences. (See, e.g., Triglia, T. et al. (1988) Nucleic Acids Res. 16:8186.) A third method, capture PCR, involves PCR amplification of DNA fragments adjacent to known sequences in human and yeast artificial chromosome DNA. (See, e.g., Lagerstrom, M. et al. (1991) PCR Methods Applic. 1:111-119.) In this method, multiple restriction enzyme digestions and ligations may be used to insert an engineered double-stranded sequence into a region of unknown sequence before performing PCR. Other methods which may be used to retrieve unknown sequences are known in the art. (See, e.g., Parker, J.D. et al. (1991) Nucleic Acids Res. 19:3055-3060). Additionally, one may use PCR, nested primers, and PROMOTERFINDER libraries (Clontech, Palo Alto CA) to walk genomic DNA. This procedure avoids the need to screen libraries and is useful in finding intron/exon junctions. For all PCR-based methods, primers may be designed using commercially available software, such as OLIGO 4.06 Primer Analysis software (National Biosciences, Plymouth MN) or another appropriate program, to be about 22 to 30 nucleotides in length, to have a

GC content of about 50% or more, and to anneal to the template at temperatures of about 68°C to 72°C.

When screening for full-length cDNAs, it is preferable to use libraries that have been size-selected to include larger cDNAs. In addition, random-primed libraries, which often include sequences containing the 5' regions of genes, are preferable for situations in which an oligo d(T) library does not yield a full-length cDNA. Genomic libraries may be useful for extension of sequence into 5' non-transcribed regulatory regions.

Capillary electrophoresis systems which are commercially available may be used to analyze the size or confirm the nucleotide sequence of sequencing or PCR products. In particular, capillary sequencing may employ flowable polymers for electrophoretic separation, four different nucleotide-specific, laser-stimulated fluorescent dyes, and a charge coupled device camera for detection of the emitted wavelengths. Output/light intensity may be converted to electrical signal using appropriate software (e.g., GENOTYPER and SEQUENCE NAVIGATOR, PE Biosystems), and the entire process from loading of samples to computer analysis and electronic data display may be computer controlled. Capillary electrophoresis is especially preferable for sequencing small DNA fragments which may be present in limited amounts in a particular sample.

In another embodiment of the invention, polynucleotide sequences or fragments thereof which encode GBAP may be cloned in recombinant DNA molecules that direct expression of GBAP, or fragments or functional equivalents thereof, in appropriate host cells. Due to the inherent degeneracy of the genetic code, other DNA sequences which encode substantially the same or a functionally equivalent amino acid sequence may be produced and used to express GBAP.

The nucleotide sequences of the present invention can be engineered using methods generally known in the art in order to alter GBAP-encoding sequences for a variety of purposes including, but not limited to, modification of the cloning, processing, and/or expression of the gene product. DNA shuffling by random fragmentation and PCR reassembly of gene fragments and synthetic oligonucleotides may be used to engineer the nucleotide sequences. For example, oligonucleotide-mediated site-directed mutagenesis may be used to introduce mutations that create new restriction sites, alter glycosylation patterns, change codon preference, produce splice variants, and so forth.

The nucleotides of the present invention may be subjected to DNA shuffling techniques such as MOLECULARBREEDING (Maxygen Inc., Santa Clara CA; described in U.S. Patent Number 5,837,458; Chang, C.-C. et al. (1999) Nat. Biotechnol. 17:793-797; Christians, F.C. et al. (1999) Nat. Biotechnol. 17:259-264; and Crameri, A. et al. (1996) Nat. Biotechnol. 14:315-319) to alter or improve the biological properties of GBAP, such as its biological or enzymatic activity or its ability to bind to other molecules or compounds. DNA shuffling is a process by which a library of gene variants is produced using PCR-mediated recombination of gene fragments. The library is then

subjected to selection or screening procedures that identify those gene variants with the desired properties. These preferred variants may then be pooled and further subjected to recursive rounds of DNA shuffling and selection/screening. Thus, genetic diversity is created through "artificial" breeding and rapid molecular evolution. For example, fragments of a single gene containing random point mutations may be recombined, screened, and then reshuffled until the desired properties are optimized. Alternatively, fragments of a given gene may be recombined with fragments of homologous genes in the same gene family, either from the same or different species, thereby maximizing the genetic diversity of multiple naturally occurring genes in a directed and controllable manner.

In another embodiment, sequences encoding GBAP may be synthesized, in whole or in part, using chemical methods well known in the art. (See, e.g., Caruthers, M.H. et al. (1980) Nucleic Acids Symp. Ser. 7:215-223; and Horn, T. et al. (1980) Nucleic Acids Symp. Ser. 7:225-232.) Alternatively, GBAP itself or a fragment thereof may be synthesized using chemical methods. For example, peptide synthesis can be performed using various solution-phase or solid-phase techniques. (See, e.g., Creighton, T. (1984) Proteins, Structures and Molecular Properties, WH Freeman, New York NY, pp. 55-60; and Roberge, J.Y. et al. (1995) Science 269:202-204.) Automated synthesis may be achieved using the ABI 431A peptide synthesizer (PE Biosystems). Additionally, the amino acid sequence of GBAP, or any part thereof, may be altered during direct synthesis and/or combined with sequences from other proteins, or any part thereof, to produce a variant polypeptide or a polypeptide having a sequence of a naturally occurring polypeptide.

The peptide may be substantially purified by preparative high performance liquid chromatography. (See, e.g., Chiez, R.M. and F.Z. Regnier (1990) Methods Enzymol. 182:392-421.) The composition of the synthetic peptides may be confirmed by amino acid analysis or by sequencing. (See, e.g., Creighton, supra, pp. 28-53.)

In order to express a biologically active GBAP, the nucleotide sequences encoding GBAP or derivatives thereof may be inserted into an appropriate expression vector, i.e., a vector which contains the necessary elements for transcriptional and translational control of the inserted coding sequence in a suitable host. These elements include regulatory sequences, such as enhancers, constitutive and inducible promoters, and 5' and 3' untranslated regions in the vector and in polynucleotide sequences encoding GBAP. Such elements may vary in their strength and specificity. Specific initiation signals may also be used to achieve more efficient translation of sequences encoding GBAP. Such signals include the ATG initiation codon and adjacent sequences, e.g. the Kozak sequence. In cases where sequences encoding GBAP and its initiation codon and upstream regulatory sequences are inserted into the appropriate expression vector, no additional transcriptional or translational control signals may be needed. However, in cases where only coding sequence, or a fragment thereof, is inserted, exogenous

translational control signals including an in-frame ATG initiation codon should be provided by the vector. Exogenous translational elements and initiation codons may be of various origins, both natural and synthetic. The efficiency of expression may be enhanced by the inclusion of enhancers appropriate for the particular host cell system used. (See, e.g., Scharf, D. et al. (1994) Results Probl. Cell Differ. 20:125-162.)

Methods which are well known to those skilled in the art may be used to construct expression vectors containing sequences encoding GBAP and appropriate transcriptional and translational control elements. These methods include in vitro recombinant DNA techniques, synthetic techniques, and in vivo genetic recombination. (See, e.g., Sambrook, J. et al. (1989) Molecular Cloning, A Laboratory Manual, Cold Spring Harbor Press, Plainview NY, ch. 4, 8, and 16-17; Ausubel, F.M. et al. (1995) Current Protocols in Molecular Biology, John Wiley & Sons, New York NY, ch. 9, 13, and 16.)

A variety of expression vector/host systems may be utilized to contain and express sequences encoding GBAP. These include, but are not limited to, microorganisms such as bacteria transformed with recombinant bacteriophage, plasmid, or cosmid DNA expression vectors; yeast transformed with yeast expression vectors; insect cell systems infected with viral expression vectors (e.g., baculovirus); plant cell systems transformed with viral expression vectors (e.g., cauliflower mosaic virus, CaMV, or tobacco mosaic virus, TMV) or with bacterial expression vectors (e.g., Ti or pBR322 plasmids); or animal cell systems. (See, e.g., Sambrook, supra; Ausubel, supra; Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509; Bitter, G.A. et al. (1987) Methods Enzymol. 153:516-544; Scorer, C.A. et al. (1994) Bio/Technology 12:181-184; Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945; Takamatsu, N. (1987) EMBO J. 6:307-311; Coruzzi, G. et al. (1984) EMBO J. 3:1671-1680; Broglie, R. et al. (1984) Science 224:838-843; Winter, J. et al. (1991) Results Probl. Cell Differ. 17:85-105; The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196; Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659; and Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.) Expression vectors derived from retroviruses, adenoviruses, or herpes or vaccinia viruses, or from various bacterial plasmids, may be used for delivery of nucleotide sequences to the targeted organ, tissue, or cell population. (See, e.g., Di Nicola, M. et al. (1998) Cancer Gen. Ther. 5(6):350-356; Yu, M. et al., (1993) Proc. Natl. Acad. Sci. USA 90(13):6340-6344; Buller, R.M. et al. (1985) Nature 317(6040):813-815; McGregor, D.P. et al. (1994) Mol. Immunol. 31(3):219-226; and Verma, I.M. and N. Somia (1997) Nature 389:239-242.) The invention is not limited by the host cell employed.

In bacterial systems, a number of cloning and expression vectors may be selected depending upon the use intended for polynucleotide sequences encoding GBAP. For example, routine cloning, subcloning, and propagation of polynucleotide sequences encoding GBAP can be achieved using a

multifunctional E. coli vector such as PBLUESCRIPT (Stratagene, La Jolla CA) or PSPORT1 plasmid (Life Technologies). Ligation of sequences encoding GBAP into the vector's multiple cloning site disrupts the *lacZ* gene, allowing a colorimetric screening procedure for identification of transformed bacteria containing recombinant molecules. In addition, these vectors may be useful for in vitro transcription, dideoxy sequencing, single strand rescue with helper phage, and creation of nested deletions in the cloned sequence. (See, e.g., Van Heeke, G. and S.M. Schuster (1989) J. Biol. Chem. 264:5503-5509.) When large quantities of GBAP are needed, e.g. for the production of antibodies, vectors which direct high level expression of GBAP may be used. For example, vectors containing the strong, inducible T5 or T7 bacteriophage promoter may be used.

Yeast expression systems may be used for production of GBAP. A number of vectors containing constitutive or inducible promoters, such as alpha factor, alcohol oxidase, and PGH promoters, may be used in the yeast Saccharomyces cerevisiae or Pichia pastoris. In addition, such vectors direct either the secretion or intracellular retention of expressed proteins and enable integration of foreign sequences into the host genome for stable propagation. (See, e.g., Ausubel, 1995, supra; Bitter, supra; and Scorer, supra.)

Plant systems may also be used for expression of GBAP. Transcription of sequences encoding GBAP may be driven viral promoters, e.g., the 35S and 19S promoters of CaMV used alone or in combination with the omega leader sequence from TMV (Takamatsu, N. (1987) EMBO J. 6:307-311). Alternatively, plant promoters such as the small subunit of RUBISCO or heat shock promoters may be used. (See, e.g., Coruzzi, supra; Broglie, supra; and Winter, supra.) These constructs can be introduced into plant cells by direct DNA transformation or pathogen-mediated transfection. (See, e.g., The McGraw Hill Yearbook of Science and Technology (1992) McGraw Hill, New York NY, pp. 191-196.)

In mammalian cells, a number of viral-based expression systems may be utilized. In cases where an adenovirus is used as an expression vector, sequences encoding GBAP may be ligated into an adenovirus transcription/translation complex consisting of the late promoter and tripartite leader sequence. Insertion in a non-essential E1 or E3 region of the viral genome may be used to obtain infective virus which expresses GBAP in host cells. (See, e.g., Logan, J. and T. Shenk (1984) Proc. Natl. Acad. Sci. USA 81:3655-3659.) In addition, transcription enhancers, such as the Rous sarcoma virus (RSV) enhancer, may be used to increase expression in mammalian host cells. SV40 or EBV-based vectors may also be used for high-level protein expression.

Human artificial chromosomes (HACs) may also be employed to deliver larger fragments of DNA than can be contained in and expressed from a plasmid. HACs of about 6 kb to 10 Mb are constructed and delivered via conventional delivery methods (liposomes, polycationic amino polymers, or vesicles) for therapeutic purposes. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355.)

For long term production of recombinant proteins in mammalian systems, stable expression of GBAP in cell lines is preferred. For example, sequences encoding GBAP can be transformed into cell lines using expression vectors which may contain viral origins of replication and/or endogenous expression elements and a selectable marker gene on the same or on a separate vector. Following the introduction of the vector, cells may be allowed to grow for about 1 to 2 days in enriched media before being switched to selective media. The purpose of the selectable marker is to confer resistance to a selective agent, and its presence allows growth and recovery of cells which successfully express the introduced sequences. Resistant clones of stably transformed cells may be propagated using tissue culture techniques appropriate to the cell type.

Any number of selection systems may be used to recover transformed cell lines. These include, but are not limited to, the herpes simplex virus thymidine kinase and adenine phosphoribosyltransferase genes, for use in *tk*⁻ and *apr*⁻ cells, respectively. (See, e.g., Wigler, M. et al. (1977) Cell 11:223-232; Lowy, I. et al. (1980) Cell 22:817-823.) Also, antimetabolite, antibiotic, or herbicide resistance can be used as the basis for selection. For example, *dhfr* confers resistance to methotrexate; *neo* confers resistance to the aminoglycosides neomycin and G-418; and *als* and *pat* confer resistance to chlorsulfuron and phosphinotricin acetyltransferase, respectively. (See, e.g., Wigler, M. et al. (1980) Proc. Natl. Acad. Sci. USA 77:3567-3570; Colbere-Garapin, F. et al. (1981) J. Mol. Biol. 150:1-14.) Additional selectable genes have been described, e.g., *trpB* and *hisD*, which alter cellular requirements for metabolites. (See, e.g., Hartman, S.C. and R.C. Mulligan (1988) Proc. Natl. Acad. Sci. USA 85:8047-8051.) Visible markers, e.g., anthocyanins, green fluorescent proteins (GFP; Clontech), ß glucuronidase and its substrate ß-glucuronide, or luciferase and its substrate luciferin may be used. These markers can be used not only to identify transformants, but also to quantify the amount of transient or stable protein expression attributable to a specific vector system. (See, e.g., Rhodes, C.A. (1995) Methods Mol. Biol. 55:121-131.)

Although the presence/absence of marker gene expression suggests that the gene of interest is also present, the presence and expression of the gene may need to be confirmed. For example, if the sequence encoding GBAP is inserted within a marker gene sequence, transformed cells containing sequences encoding GBAP can be identified by the absence of marker gene function. Alternatively, a marker gene can be placed in tandem with a sequence encoding GBAP under the control of a single promoter. Expression of the marker gene in response to induction or selection usually indicates expression of the tandem gene as well.

In general, host cells that contain the nucleic acid sequence encoding GBAP and that express GBAP may be identified by a variety of procedures known to those of skill in the art. These procedures include, but are not limited to, DNA-DNA or DNA-RNA hybridizations, PCR amplification, and protein bioassay or immunoassay techniques which include membrane, solution, or chip based

technologies for the detection and/or quantification of nucleic acid or protein sequences.

Immunological methods for detecting and measuring the expression of GBAP using either specific polyclonal or monoclonal antibodies are known in the art. Examples of such techniques include enzyme-linked immunosorbent assays (ELISAs), radioimmunoassays (RIAs), and fluorescence activated cell sorting (FACS). A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering epitopes on GBAP is preferred, but a competitive binding assay may be employed. These and other assays are well known in the art. (See, e.g., Hampton, R. et al. (1990) Serological Methods, a Laboratory Manual, APS Press, St. Paul MN, Sect. IV; Coligan, J.E. et al. (1997) Current Protocols in Immunology, Greene Pub. Associates and Wiley-Interscience, New York NY; and Pound, J.D. (1998) Immunochemical Protocols, Humana Press, Totowa NJ.)

A wide variety of labels and conjugation techniques are known by those skilled in the art and may be used in various nucleic acid and amino acid assays. Means for producing labeled hybridization or PCR probes for detecting sequences related to polynucleotides encoding GBAP include oligolabeling, nick translation, end-labeling, or PCR amplification using a labeled nucleotide. Alternatively, the sequences encoding GBAP, or any fragments thereof, may be cloned into a vector for the production of an mRNA probe. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by addition of an appropriate RNA polymerase such as T7, T3, or SP6 and labeled nucleotides. These procedures may be conducted using a variety of commercially available kits, such as those provided by Amersham Pharmacia Biotech, Promega (Madison WI), and US Biochemical. Suitable reporter molecules or labels which may be used for ease of detection include radionuclides, enzymes, fluorescent, chemiluminescent, or chromogenic agents, as well as substrates, cofactors, inhibitors, magnetic particles, and the like.

Host cells transformed with nucleotide sequences encoding GBAP may be cultured under conditions suitable for the expression and recovery of the protein from cell culture. The protein produced by a transformed cell may be secreted or retained intracellularly depending on the sequence and/or the vector used. As will be understood by those of skill in the art, expression vectors containing polynucleotides which encode GBAP may be designed to contain signal sequences which direct secretion of GBAP through a prokaryotic or eukaryotic cell membrane.

In addition, a host cell strain may be chosen for its ability to modulate expression of the inserted sequences or to process the expressed protein in the desired fashion. Such modifications of the polypeptide include, but are not limited to, acetylation, carboxylation, glycosylation, phosphorylation, lipidation, and acylation. Post-translational processing which cleaves a "prepro" or "pro" form of the protein may also be used to specify protein targeting, folding, and/or activity. Different host cells which have specific cellular machinery and characteristic mechanisms for post-translational activities (e.g., CHO, HeLa, MDCK, HEK293, and WI38) are available from the American Type Culture

Collection (ATCC, Manassas VA) and may be chosen to ensure the correct modification and processing of the foreign protein.

In another embodiment of the invention, natural, modified, or recombinant nucleic acid sequences encoding GBAP may be ligated to a heterologous sequence resulting in translation of a fusion protein in any of the aforementioned host systems. For example, a chimeric GBAP protein containing a heterologous moiety that can be recognized by a commercially available antibody may facilitate the screening of peptide libraries for inhibitors of GBAP activity. Heterologous protein and peptide moieties may also facilitate purification of fusion proteins using commercially available affinity matrices. Such moieties include, but are not limited to, glutathione S-transferase (GST), maltose binding protein (MBP), thioredoxin (Trx), calmodulin binding peptide (CBP), 6-His, FLAG, *c-myc*, and hemagglutinin (HA). GST, MBP, Trx, CBP, and 6-His enable purification of their cognate fusion proteins on immobilized glutathione, maltose, phenylarsine oxide, calmodulin, and metal-chelate resins, respectively. FLAG, *c-myc*, and hemagglutinin (HA) enable immunoaffinity purification of fusion proteins using commercially available monoclonal and polyclonal antibodies that specifically recognize these epitope tags. A fusion protein may also be engineered to contain a proteolytic cleavage site located between the GBAP encoding sequence and the heterologous protein sequence, so that GBAP may be cleaved away from the heterologous moiety following purification. Methods for fusion protein expression and purification are discussed in Ausubel (1995, supra, ch. 10). A variety of commercially available kits may also be used to facilitate expression and purification of fusion proteins.

In a further embodiment of the invention, synthesis of radiolabeled GBAP may be achieved in vitro using the TNT rabbit reticulocyte lysate or wheat germ extract system (Promega). These systems couple transcription and translation of protein-coding sequences operably associated with the T7, T3, or SP6 promoters. Translation takes place in the presence of a radiolabeled amino acid precursor, for example, ^{35}S-methionine.

GBAP of the present invention or fragments thereof may be used to screen for compounds that specifically bind to GBAP. At least one and up to a plurality of test compounds may be screened for specific binding to GBAP. Examples of test compounds include antibodies, oligonucleotides, proteins (e.g., receptors), or small molecules.

In one embodiment, the compound thus identified is closely related to the natural ligand of GBAP, e.g., a ligand or fragment thereof, a natural substrate, a structural or functional mimetic, or a natural binding partner. (See, Coligan, J.E. et al. (1991) Current Protocols in Immunology 1(2): Chapter 5.) Similarly, the compound can be closely related to the natural receptor to which GBAP binds, or to at least a fragment of the receptor, e.g., the ligand binding site. In either case, the compound can be rationally designed using known techniques. In one embodiment, screening for these compounds involves producing appropriate cells which express GBAP, either as a secreted

protein or on the cell membrane. Preferred cells include cells from mammals, yeast, Drosophila, or E. coli. Cells expressing GBAP or cell membrane fractions which contain GBAP are then contacted with a test compound and binding, stimulation, or inhibition of activity of either GBAP or the compound is analyzed.

An assay may simply test binding of a test compound to the polypeptide, wherein binding is detected by a fluorophore, radioisotope, enzyme conjugate, or other detectable label. For example, the assay may comprise the steps of combining at least one test compound with GBAP, either in solution or affixed to a solid support, and detecting the binding of GBAP to the compound. Alternatively, the assay may detect or measure binding of a test compound in the presence of a labeled competitor. Additionally, the assay may be carried out using cell-free preparations, chemical libraries, or natural product mixtures, and the test compound(s) may be free in solution or affixed to a solid support.

GBAP of the present invention or fragments thereof may be used to screen for compounds that modulate the activity of GBAP. Such compounds may include agonists, antagonists, or partial or inverse agonists. In one embodiment, an assay is performed under conditions permissive for GBAP activity, wherein GBAP is combined with at least one test compound, and the activity of GBAP in the presence of a test compound is compared with the activity of GBAP in the absence of the test compound. A change in the activity of GBAP in the presence of the test compound is indicative of a compound that modulates the activity of GBAP. Alternatively, a test compound is combined with an in vitro or cell-free system comprising GBAP under conditions suitable for GBAP activity, and the assay is performed. In either of these assays, a test compound which modulates the activity of GBAP may do so indirectly and need not come in direct contact with the test compound. At least one and up to a plurality of test compounds may be screened.

In another embodiment, polynucleotides encoding GBAP or their mammalian homologs may be "knocked out" in an animal model system using homologous recombination in embryonic stem (ES) cells. Such techniques are well known in the art and are useful for the generation of animal models of human disease. (See, e.g., U.S. Patent No. 5,175,383 and U.S. Patent No. 5,767,337.) For example, mouse ES cells, such as the mouse 129/SvJ cell line, are derived from the early mouse embryo and grown in culture. The ES cells are transformed with a vector containing the gene of interest disrupted by a marker gene, e.g., the neomycin phosphotransferase gene (neo; Capecchi, M.R. (1989) Science 244:1288-1292). The vector integrates into the corresponding region of the host genome by homologous recombination. Alternatively, homologous recombination takes place using the Cre-loxP system to knockout a gene of interest in a tissue- or developmental stage-specific manner (Marth, J.D. (1996) Clin. Invest. 97:1999-2002; Wagner, K.U. et al. (1997) Nucleic Acids Res. 25:4323-4330). Transformed ES cells are identified and microinjected into mouse cell blastocysts such as those from the C57BL/6 mouse strain. The blastocysts are surgically transferred

to pseudopregnant dams, and the resulting chimeric progeny are genotyped and bred to produce heterozygous or homozygous strains. Transgenic animals thus generated may be tested with potential therapeutic or toxic agents.

Polynucleotides encoding GBAP may also be manipulated in vitro in ES cells derived from human blastocysts. Human ES cells have the potential to differentiate into at least eight separate cell lineages including endoderm, mesoderm, and ectodermal cell types. These cell lineages differentiate into, for example, neural cells, hematopoietic lineages, and cardiomyocytes (Thomson, J.A. et al. (1998) Science 282:1145-1147).

Polynucleotides encoding GBAP can also be used to create "knockin" humanized animals (pigs) or transgenic animals (mice or rats) to model human disease. With knockin technology, a region of a polynucleotide encoding GBAP is injected into animal ES cells, and the injected sequence integrates into the animal cell genome. Transformed cells are injected into blastulae, and the blastulae are implanted as described above. Transgenic progeny or inbred lines are studied and treated with potential pharmaceutical agents to obtain information on treatment of a human disease. Alternatively, a mammal inbred to overexpress GBAP, e.g., by secreting GBAP in its milk, may also serve as a convenient source of that protein (Janne, J. et al. (1998) Biotechnol. Annu. Rev. 4:55-74).

THERAPEUTICS

Chemical and structural similarity, e.g., in the context of sequences and motifs, exists between regions of GBAP and GTP-binding associated proteins. In addition, the expression of GBAP is closely associated with reproductive tissues, inflammation and the immune response, trauma, cell proliferation, and cancer. Therefore, GBAP appears to play a role in immune system, reproductive, nervous system, and cell signaling disorders, and cell proliferative disorders including cancer. In the treatment of disorders associated with increased GBAP expression or activity, it is desirable to decrease the expression or activity of GBAP. In the treatment of disorders associated with decreased GBAP expression or activity, it is desirable to increase the expression or activity of GBAP.

Therefore, in one embodiment, GBAP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GBAP. Examples of such disorders include, but are not limited to, an immune system disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable

bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis, cancer of the breast, fibrocystic breast disease, and galactorrhea, a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a nervous system disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; a cell signaling disorder including endocrine disorders such as disorders of the hypothalamus and pituitary resulting from lesions such as primary brain tumors, adenomas, infarction associated with pregnancy, hypophysectomy, aneurysms, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma; disorders associated with hyperpituitarism including acromegaly, giantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma; disorders associated with

hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection; disorders associated with hyperparathyroidism including Conn disease (chronic hypercalemia); pancreatic disorders such as Type I or Type II diabetes mellitus and associated complications; disorders associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis; disorders associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbations of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis; and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, hypergonadal disorders associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α-reductase, and gynecomastia; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus.

In another embodiment, a vector capable of expressing GBAP or a fragment or derivative thereof may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GBAP including, but not limited to, those described above.

In a further embodiment, a pharmaceutical composition comprising a substantially purified GBAP in conjunction with a suitable pharmaceutical carrier may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GBAP including, but not limited to, those provided above.

In still another embodiment, an agonist which modulates the activity of GBAP may be administered to a subject to treat or prevent a disorder associated with decreased expression or activity of GBAP including, but not limited to, those listed above.

In a further embodiment, an antagonist of GBAP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of GBAP. Examples of such disorders include, but are not limited to, those immune system, reproductive, nervous system, and cell signaling disorders, and cell proliferative disorders including cancer, described above. In one aspect, an antibody which specifically binds GBAP may be used directly as an antagonist or indirectly as a targeting or delivery mechanism for bringing a pharmaceutical agent to cells or tissues which express GBAP.

In an additional embodiment, a vector expressing the complement of the polynucleotide encoding GBAP may be administered to a subject to treat or prevent a disorder associated with increased expression or activity of GBAP including, but not limited to, those described above.

In other embodiments, any of the proteins, antagonists, antibodies, agonists, complementary sequences, or vectors of the invention may be administered in combination with other appropriate therapeutic agents. Selection of the appropriate agents for use in combination therapy may be made by one of ordinary skill in the art, according to conventional pharmaceutical principles. The combination of therapeutic agents may act synergistically to effect the treatment or prevention of the various disorders described above. Using this approach, one may be able to achieve therapeutic efficacy with lower dosages of each agent, thus reducing the potential for adverse side effects.

An antagonist of GBAP may be produced using methods which are generally known in the art. In particular, purified GBAP may be used to produce antibodies or to screen libraries of pharmaceutical agents to identify those which specifically bind GBAP. Antibodies to GBAP may also be generated using methods that are well known in the art. Such antibodies may include, but are not limited to, polyclonal, monoclonal, chimeric, and single chain antibodies, Fab fragments, and fragments produced by a Fab expression library. Neutralizing antibodies (i.e., those which inhibit dimer formation) are generally preferred for therapeutic use.

For the production of antibodies, various hosts including goats, rabbits, rats, mice, humans, and others may be immunized by injection with GBAP or with any fragment or oligopeptide thereof which has immunogenic properties. Depending on the host species, various adjuvants may be used to increase immunological response. Such adjuvants include, but are not limited to, Freund's, mineral gels such as aluminum hydroxide, and surface active substances such as lysolecithin, pluronic polyols, polyanions, peptides, oil emulsions, KLH, and dinitrophenol. Among adjuvants used in humans, BCG (bacilli Calmette-Guerin) and Corynebacterium parvum are especially preferable.

It is preferred that the oligopeptides, peptides, or fragments used to induce antibodies to GBAP have an amino acid sequence consisting of at least about 5 amino acids, and generally will consist of at least about 10 amino acids. It is also preferable that these oligopeptides, peptides, or fragments are identical to a portion of the amino acid sequence of the natural protein. Short stretches of GBAP amino acids may be fused with those of another protein, such as KLH, and antibodies to the chimeric molecule may be produced.

Monoclonal antibodies to GBAP may be prepared using any technique which provides for the production of antibody molecules by continuous cell lines in culture. These include, but are not limited to, the hybridoma technique, the human B-cell hybridoma technique, and the EBV-hybridoma technique. (See, e.g., Kohler, G. et al. (1975) Nature 256:495-497; Kozbor, D. et al. (1985) J. Immunol. Methods 81:31-42; Cote, R.J. et al. (1983) Proc. Natl. Acad. Sci. USA 80:2026-2030; and

Cole, S.P. et al. (1984) Mol. Cell Biol. 62:109-120.)

In addition, techniques developed for the production of "chimeric antibodies," such as the splicing of mouse antibody genes to human antibody genes to obtain a molecule with appropriate antigen specificity and biological activity, can be used. (See, e.g., Morrison, S.L. et al. (1984) Proc. Natl. Acad. Sci. USA 81:6851-6855; Neuberger, M.S. et al. (1984) Nature 312:604-608; and Takeda, S. et al. (1985) Nature 314:452-454.) Alternatively, techniques described for the production of single chain antibodies may be adapted, using methods known in the art, to produce GBAP-specific single chain antibodies. Antibodies with related specificity, but of distinct idiotypic composition, may be generated by chain shuffling from random combinatorial immunoglobulin libraries. (See, e.g., Burton, D.R. (1991) Proc. Natl. Acad. Sci. USA 88:10134-10137.)

Antibodies may also be produced by inducing in vivo production in the lymphocyte population or by screening immunoglobulin libraries or panels of highly specific binding reagents as disclosed in the literature. (See, e.g., Orlandi, R. et al. (1989) Proc. Natl. Acad. Sci. USA 86:3833-3837; Winter, G. et al. (1991) Nature 349:293-299.)

Antibody fragments which contain specific binding sites for GBAP may also be generated. For example, such fragments include, but are not limited to, $F(ab')_2$ fragments produced by pepsin digestion of the antibody molecule and Fab fragments generated by reducing the disulfide bridges of the F(ab')2 fragments. Alternatively, Fab expression libraries may be constructed to allow rapid and easy identification of monoclonal Fab fragments with the desired specificity. (See, e.g., Huse, W.D. et al. (1989) Science 246:1275-1281.)

Various immunoassays may be used for screening to identify antibodies having the desired specificity. Numerous protocols for competitive binding or immunoradiometric assays using either polyclonal or monoclonal antibodies with established specificities are well known in the art. Such immunoassays typically involve the measurement of complex formation between GBAP and its specific antibody. A two-site, monoclonal-based immunoassay utilizing monoclonal antibodies reactive to two non-interfering GBAP epitopes is generally used, but a competitive binding assay may also be employed (Pound, supra).

Various methods such as Scatchard analysis in conjunction with radioimmunoassay techniques may be used to assess the affinity of antibodies for GBAP. Affinity is expressed as an association constant, K_a, which is defined as the molar concentration of GBAP-antibody complex divided by the molar concentrations of free antigen and free antibody under equilibrium conditions. The K_a determined for a preparation of polyclonal antibodies, which are heterogeneous in their affinities for multiple GBAP epitopes, represents the average affinity, or avidity, of the antibodies for GBAP. The K_a determined for a preparation of monoclonal antibodies, which are monospecific for a particular GBAP epitope, represents a true measure of affinity. High-affinity antibody preparations with K_a ranging from

about 10^9 to 10^{12} L/mole are preferred for use in immunoassays in which the GBAP-antibody complex must withstand rigorous manipulations. Low-affinity antibody preparations with K_a ranging from about 10^6 to 10^7 L/mole are preferred for use in immunopurification and similar procedures which ultimately require dissociation of GBAP, preferably in active form, from the antibody (Catty, D. (1988) Antibodies, Volume I: A Practical Approach, IRL Press, Washington DC; Liddell, J.E. and A. Cryer (1991) A Practical Guide to Monoclonal Antibodies, John Wiley & Sons, New York NY).

The titer and avidity of polyclonal antibody preparations may be further evaluated to determine the quality and suitability of such preparations for certain downstream applications. For example, a polyclonal antibody preparation containing at least 1-2 mg specific antibody/ml, preferably 5-10 mg specific antibody/ml, is generally employed in procedures requiring precipitation of GBAP-antibody complexes. Procedures for evaluating antibody specificity, titer, and avidity, and guidelines for antibody quality and usage in various applications, are generally available. (See, e.g., Catty, supra, and Coligan et al., supra.)

In another embodiment of the invention, the polynucleotides encoding GBAP, or any fragment or complement thereof, may be used for therapeutic purposes. In one aspect, modifications of gene expression can be achieved by designing complementary sequences or antisense molecules (DNA, RNA, PNA, or modified oligonucleotides) to the coding or regulatory regions of the gene encoding GBAP. Such technology is well known in the art, and antisense oligonucleotides or larger fragments can be designed from various locations along the coding or control regions of sequences encoding GBAP. (See, e.g., Agrawal, S., ed. (1996) Antisense Therapeutics, Humana Press Inc., Totawa NJ.)

In therapeutic use, any gene delivery system suitable for introduction of the antisense sequences into appropriate target cells can be used. Antisense sequences can be delivered intracellularly in the form of an expression plasmid which, upon transcription, produces a sequence complementary to at least a portion of the cellular sequence encoding the target protein. (See, e.g., Slater, J.E. et al. (1998) J. Allergy Clin. Immunol. 102(3):469-475; and Scanlon, K.J. et al. (1995) 9(13):1288-1296.) Antisense sequences can also be introduced intracellularly through the use of viral vectors, such as retrovirus and adeno-associated virus vectors. (See, e.g., Miller, A.D. (1990) Blood 76:271; Ausubel, supra; Uckert, W. and W. Walther (1994) Pharmacol. Ther. 63(3):323-347.) Other gene delivery mechanisms include liposome-derived systems, artificial viral envelopes, and other systems known in the art. (See, e.g., Rossi, J.J. (1995) Br. Med. Bull. 51(1):217-225; Boado, R.J. et al. (1998) J. Pharm. Sci. 87(11):1308-1315; and Morris, M.C. et al. (1997) Nucleic Acids Res. 25(14):2730-2736.)

In another embodiment of the invention, polynucleotides encoding GBAP may be used for somatic or germline gene therapy. Gene therapy may be performed to (i) correct a genetic deficiency (e.g., in the cases of severe combined immunodeficiency (SCID)-X1 disease characterized by X-linked

inheritance (Cavazzana-Calvo, M. et al. (2000) Science 288:669-672), severe combined immunodeficiency syndrome associated with an inherited adenosine deaminase (ADA) deficiency (Blaese, R.M. et al. (1995) Science 270:475-480; Bordignon, C. et al. (1995) Science 270:470-475), cystic fibrosis (Zabner, J. et al. (1993) Cell 75:207-216; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:643-666; Crystal, R.G. et al. (1995) Hum. Gene Therapy 6:667-703), thalassamias, familial hypercholesterolemia, and hemophilia resulting from Factor VIII or Factor IX deficiencies (Crystal, R.G. (1995) Science 270:404-410; Verma, I.M. and Somia, N. (1997) Nature 389:239-242)), (ii) express a conditionally lethal gene product (e.g., in the case of cancers which result from unregulated cell proliferation), or (iii) express a protein which affords protection against intracellular parasites (e.g., against human retroviruses, such as human immunodeficiency virus (HIV) (Baltimore, D. (1988) Nature 335:395-396; Poeschla, E. et al. (1996) Proc. Natl. Acad. Sci. USA. 93:11395-11399), hepatitis B or C virus (HBV, HCV); fungal parasites, such as Candida albicans and Paracoccidioides brasiliensis; and protozoan parasites such as Plasmodium falciparum and Trypanosoma cruzi). In the case where a genetic deficiency in GBAP expression or regulation causes disease, the expression of GBAP from an appropriate population of transduced cells may alleviate the clinical manifestations caused by the genetic deficiency.

In a further embodiment of the invention, diseases or disorders caused by deficiencies in GBAP are treated by constructing mammalian expression vectors encoding GBAP and introducing these vectors by mechanical means into GBAP-deficient cells. Mechanical transfer technologies for use with cells in vivo or ex vitro include (i) direct DNA microinjection into individual cells, (ii) ballistic gold particle delivery, (iii) liposome-mediated transfection, (iv) receptor-mediated gene transfer, and (v) the use of DNA transposons (Morgan, R.A. and W.F. Anderson (1993) Annu. Rev. Biochem. 62:191-217; Ivics, Z. (1997) Cell 91:501-510; Boulay, J-L. and H. Récipon (1998) Curr. Opin. Biotechnol. 9:445-450).

Expression vectors that may be effective for the expression of GBAP include, but are not limited to, the PCDNA 3.1, EPITAG, PRCCMV2, PREP, PVAX vectors (Invitrogen, Carlsbad CA), PCMV-SCRIPT, PCMV-TAG, PEGSH/PERV (Stratagene, La Jolla CA), and PTET-OFF, PTET-ON, PTRE2, PTRE2-LUC, PTK-HYG (Clontech, Palo Alto CA). GBAP may be expressed using (i) a constitutively active promoter, (e.g., from cytomegalovirus (CMV), Rous sarcoma virus (RSV), SV40 virus, thymidine kinase (TK), or β-actin genes), (ii) an inducible promoter (e.g., the tetracycline-regulated promoter (Gossen, M. and H. Bujard (1992) Proc. Natl. Acad. Sci. USA 89:5547-5551; Gossen, M. et al. (1995) Science 268:1766-1769; Rossi, F.M.V. and H.M. Blau (1998) Curr. Opin. Biotechnol. 9:451-456), commercially available in the T-REX plasmid (Invitrogen)); the ecdysone-inducible promoter (available in the plasmids PVGRXR and PIND; Invitrogen); the FK506/rapamycin inducible promoter; or the RU486/mifepristone inducible promoter (Rossi, F.M.V.

and H.M. Blau, supra)), or (iii) a tissue-specific promoter or the native promoter of the endogenous gene encoding GBAP from a normal individual.

Commercially available liposome transformation kits (e.g., the PERFECT LIPID TRANSFECTION KIT, available from Invitrogen) allow one with ordinary skill in the art to deliver polynucleotides to target cells in culture and require minimal effort to optimize experimental parameters. In the alternative, transformation is performed using the calcium phosphate method (Graham, F.L. and A.J. Eb (1973) Virology 52:456-467), or by electroporation (Neumann, E. et al. (1982) EMBO J. 1:841-845). The introduction of DNA to primary cells requires modification of these standardized mammalian transfection protocols.

In another embodiment of the invention, diseases or disorders caused by genetic defects with respect to GBAP expression are treated by constructing a retrovirus vector consisting of (i) the polynucleotide encoding GBAP under the control of an independent promoter or the retrovirus long terminal repeat (LTR) promoter, (ii) appropriate RNA packaging signals, and (iii) a Rev-responsive element (RRE) along with additional retrovirus *cis*-acting RNA sequences and coding sequences required for efficient vector propagation. Retrovirus vectors (e.g., PFB and PFBNEO) are commercially available (Stratagene) and are based on published data (Riviere, I. et al. (1995) Proc. Natl. Acad. Sci. USA 92:6733-6737), incorporated by reference herein. The vector is propagated in an appropriate vector producing cell line (VPCL) that expresses an envelope gene with a tropism for receptors on the target cells or a promiscuous envelope protein such as VSVg (Armentano, D. et al. (1987) J. Virol. 61:1647-1650; Bender, M.A. et al. (1987) J. Virol. 61:1639-1646; Adam, M.A. and A.D. Miller (1988) J. Virol. 62:3802-3806; Dull, T. et al. (1998) J. Virol. 72:8463-8471; Zufferey, R. et al. (1998) J. Virol. 72:9873-9880). U.S. Patent Number 5,910,434 to Rigg ("Method for obtaining retrovirus packaging cell lines producing high transducing efficiency retroviral supernatant") discloses a method for obtaining retrovirus packaging cell lines and is hereby incorporated by reference. Propagation of retrovirus vectors, transduction of a population of cells (e.g., $CD4^+$ T-cells), and the return of transduced cells to a patient are procedures well known to persons skilled in the art of gene therapy and have been well documented (Ranga, U. et al. (1997) J. Virol. 71:7020-7029; Bauer, G. et al. (1997) Blood 89:2259-2267; Bonyhadi, M.L. (1997) J. Virol. 71:4707-4716; Ranga, U. et al. (1998) Proc. Natl. Acad. Sci. USA 95:1201-1206; Su, L. (1997) Blood 89:2283-2290).

In the alternative, an adenovirus-based gene therapy delivery system is used to deliver polynucleotides encoding GBAP to cells which have one or more genetic abnormalities with respect to the expression of GBAP. The construction and packaging of adenovirus-based vectors are well known to those with ordinary skill in the art. Replication defective adenovirus vectors have proven to be versatile for importing genes encoding immunoregulatory proteins into intact islets in the pancreas (Csete, M.E. et al. (1995) Transplantation 27:263-268). Potentially useful adenoviral vectors are

described in U.S. Patent Number 5,707,618 to Armentano ("Adenovirus vectors for gene therapy"), hereby incorporated by reference. For adenoviral vectors, see also Antinozzi, P.A. et al. (1999) Annu. Rev. Nutr. 19:511-544; and Verma, I.M. and N. Somia (1997) Nature 18:389:239-242, both incorporated by reference herein.

In another alternative, a herpes-based, gene therapy delivery system is used to deliver polynucleotides encoding GBAP to target cells which have one or more genetic abnormalities with respect to the expression of GBAP. The use of herpes simplex virus (HSV)-based vectors may be especially valuable for introducing GBAP to cells of the central nervous system, for which HSV has a tropism. The construction and packaging of herpes-based vectors are well known to those with ordinary skill in the art. A replication-competent herpes simplex virus (HSV) type 1-based vector has been used to deliver a reporter gene to the eyes of primates (Liu, X. et al. (1999) Exp. Eye Res.169:385-395). The construction of a HSV-1 virus vector has also been disclosed in detail in U.S. Patent Number 5,804,413 to DeLuca ("Herpes simplex virus strains for gene transfer"), which is hereby incorporated by reference. U.S. Patent Number 5,804,413 teaches the use of recombinant HSV d92 which consists of a genome containing at least one exogenous gene to be transferred to a cell under the control of the appropriate promoter for purposes including human gene therapy. Also taught by this patent are the construction and use of recombinant HSV strains deleted for ICP4, ICP27 and ICP22. For HSV vectors, see also Goins, W.F. et al. (1999) J. Virol. 73:519-532 and Xu, H. et al. (1994) Dev. Biol. 163:152-161, hereby incorporated by reference. The manipulation of cloned herpesvirus sequences, the generation of recombinant virus following the transfection of multiple plasmids containing different segments of the large herpesvirus genomes, the growth and propagation of herpesvirus, and the infection of cells with herpesvirus are techniques well known to those of ordinary skill in the art.

In another alternative, an alphavirus (positive, single-stranded RNA virus) vector is used to deliver polynucleotides encoding GBAP to target cells. The biology of the prototypic alphavirus, Semliki Forest Virus (SFV), has been studied extensively and gene transfer vectors have been based on the SFV genome (Garoff, H. and K.-J. Li (1998) Curr. Opin. Biotech. 9:464-469). During alphavirus RNA replication, a subgenomic RNA is generated that normally encodes the viral capsid proteins. This subgenomic RNA replicates to higher levels than the full-length genomic RNA, resulting in the overproduction of capsid proteins relative to the viral proteins with enzymatic activity (e.g., protease and polymerase). Similarly, inserting the coding sequence for GBAP into the alphavirus genome in place of the capsid-coding region results in the production of a large number of GBAP-coding RNAs and the synthesis of high levels of GBAP in vector transduced cells. While alphavirus infection is typically associated with cell lysis within a few days, the ability to establish a persistent infection in hamster normal kidney cells (BHK-21) with a variant of Sindbis virus (SIN) indicates that the lytic

replication of alphaviruses can be altered to suit the needs of the gene therapy application (Dryga, S.A. et al. (1997) Virology 228:74-83). The wide host range of alphaviruses will allow the introduction of GBAP into a variety of cell types. The specific transduction of a subset of cells in a population may require the sorting of cells prior to transduction. The methods of manipulating infectious cDNA clones of alphaviruses, performing alphavirus cDNA and RNA transfections, and performing alphavirus infections, are well known to those with ordinary skill in the art.

Oligonucleotides derived from the transcription initiation site, e.g., between about positions -10 and +10 from the start site, may also be employed to inhibit gene expression. Similarly, inhibition can be achieved using triple helix base-pairing methodology. Triple helix pairing is useful because it causes inhibition of the ability of the double helix to open sufficiently for the binding of polymerases, transcription factors, or regulatory molecules. Recent therapeutic advances using triplex DNA have been described in the literature. (See, e.g., Gee, J.E. et al. (1994) in Huber, B.E. and B.I. Carr, Molecular and Immunologic Approaches, Futura Publishing, Mt. Kisco NY, pp. 163-177.) A complementary sequence or antisense molecule may also be designed to block translation of mRNA by preventing the transcript from binding to ribosomes.

Ribozymes, enzymatic RNA molecules, may also be used to catalyze the specific cleavage of RNA. The mechanism of ribozyme action involves sequence-specific hybridization of the ribozyme molecule to complementary target RNA, followed by endonucleolytic cleavage. For example, engineered hammerhead motif ribozyme molecules may specifically and efficiently catalyze endonucleolytic cleavage of sequences encoding GBAP.

Specific ribozyme cleavage sites within any potential RNA target are initially identified by scanning the target molecule for ribozyme cleavage sites, including the following sequences: GUA, GUU, and GUC. Once identified, short RNA sequences of between 15 and 20 ribonucleotides, corresponding to the region of the target gene containing the cleavage site, may be evaluated for secondary structural features which may render the oligonucleotide inoperable. The suitability of candidate targets may also be evaluated by testing accessibility to hybridization with complementary oligonucleotides using ribonuclease protection assays.

Complementary ribonucleic acid molecules and ribozymes of the invention may be prepared by any method known in the art for the synthesis of nucleic acid molecules. These include techniques for chemically synthesizing oligonucleotides such as solid phase phosphoramidite chemical synthesis. Alternatively, RNA molecules may be generated by in vitro and in vivo transcription of DNA sequences encoding GBAP. Such DNA sequences may be incorporated into a wide variety of vectors with suitable RNA polymerase promoters such as T7 or SP6. Alternatively, these cDNA constructs that synthesize complementary RNA, constitutively or inducibly, can be introduced into cell lines, cells, or tissues.

RNA molecules may be modified to increase intracellular stability and half-life. Possible modifications include, but are not limited to, the addition of flanking sequences at the 5' and/or 3' ends of the molecule, or the use of phosphorothioate or 2' O-methyl rather than phosphodiesterase linkages within the backbone of the molecule. This concept is inherent in the production of PNAs and can be extended in all of these molecules by the inclusion of nontraditional bases such as inosine, queosine, and wybutosine, as well as acetyl-, methyl-, thio-, and similarly modified forms of adenine, cytidine, guanine, thymine, and uridine which are not as easily recognized by endogenous endonucleases.

An additional embodiment of the invention encompasses a method for screening for a compound which is effective in altering expression of a polynucleotide encoding GBAP. Compounds which may be effective in altering expression of a specific polynucleotide may include, but are not limited to, oligonucleotides, antisense oligonucleotides, triple helix-forming oligonucleotides, transcription factors and other polypeptide transcriptional regulators, and non-macromolecular chemical entities which are capable of interacting with specific polynucleotide sequences. Effective compounds may alter polynucleotide expression by acting as either inhibitors or promoters of polynucleotide expression. Thus, in the treatment of disorders associated with increased GBAP expression or activity, a compound which specifically inhibits expression of the polynucleotide encoding GBAP may be therapeutically useful, and in the treament of disorders associated with decreased GBAP expression or activity, a compound which specifically promotes expression of the polynucleotide encoding GBAP may be therapeutically useful.

At least one, and up to a plurality, of test compounds may be screened for effectiveness in altering expression of a specific polynucleotide. A test compound may be obtained by any method commonly known in the art, including chemical modification of a compound known to be effective in altering polynucleotide expression; selection from an existing, commercially-available or proprietary library of naturally-occurring or non-natural chemical compounds; rational design of a compound based on chemical and/or structural properties of the target polynucleotide; and selection from a library of chemical compounds created combinatorially or randomly. A sample comprising a polynucleotide encoding GBAP is exposed to at least one test compound thus obtained. The sample may comprise, for example, an intact or permeabilized cell, or an in vitro cell-free or reconstituted biochemical system. Alterations in the expression of a polynucleotide encoding GBAP are assayed by any method commonly known in the art. Typically, the expression of a specific nucleotide is detected by hybridization with a probe having a nucleotide sequence complementary to the sequence of the polynucleotide encoding GBAP. The amount of hybridization may be quantified, thus forming the basis for a comparison of the expression of the polynucleotide both with and without exposure to one or more test compounds. Detection of a change in the expression of a polynucleotide exposed to a test compound indicates that the test compound is effective in altering the expression of

the polynucleotide. A screen for a compound effective in altering expression of a specific polynucleotide can be carried out, for example, using a Schizosaccharomyces pombe gene expression system (Atkins, D. et al. (1999) U.S. Patent No. 5,932,435; Arndt, G.M. et al. (2000) Nucleic Acids Res. 28:E15) or a human cell line such as HeLa cell (Clarke, M.L. et al. (2000) Biochem. Biophys. Res. Commun. 268:8-13). A particular embodiment of the present invention involves screening a combinatorial library of oligonucleotides (such as deoxyribonucleotides, ribonucleotides, peptide nucleic acids, and modified oligonucleotides) for antisense activity against a specific polynucleotide sequence (Bruice, T.W. et al. (1997) U.S. Patent No. 5,686,242; Bruice, T.W. et al. (2000) U.S. Patent No. 6,022,691).

Many methods for introducing vectors into cells or tissues are available and equally suitable for use in vivo, in vitro, and ex vivo. For ex vivo therapy, vectors may be introduced into stem cells taken from the patient and clonally propagated for autologous transplant back into that same patient. Delivery by transfection, by liposome injections, or by polycationic amino polymers may be achieved using methods which are well known in the art. (See, e.g., Goldman, C.K. et al. (1997) Nat. Biotechnol. 15:462-466.)

Any of the therapeutic methods described above may be applied to any subject in need of such therapy, including, for example, mammals such as humans, dogs, cats, cows, horses, rabbits, and monkeys.

An additional embodiment of the invention relates to the administration of a pharmaceutical composition which generally comprises an active ingredient formulated with a pharmaceutically acceptable excipient. Excipients may include, for example, sugars, starches, celluloses, gums, and proteins. Various formulations are commonly known and are thoroughly discussed in the latest edition of Remington's Pharmaceutical Sciences (Maack Publishing, Easton PA). Such pharmaceutical compositions may consist of GBAP, antibodies to GBAP, and mimetics, agonists, antagonists, or inhibitors of GBAP.

The pharmaceutical compositions utilized in this invention may be administered by any number of routes including, but not limited to, oral, intravenous, intramuscular, intra-arterial, intramedullary, intrathecal, intraventricular, pulmonary, transdermal, subcutaneous, intraperitoneal, intranasal, enteral, topical, sublingual, or rectal means.

Pharmaceutical compositions for pulmonary administration may be prepared in liquid or dry powder form. These compositions are generally aerosolized immediately prior to inhalation by the patient. In the case of small molecules (e.g. traditional low molecular weight organic drugs), aerosol delivery of fast-acting formulations is well-known in the art. In the case of macromolecules (e.g. larger peptides and proteins), recent developments in the field of pulmonary delivery via the alveolar region of the lung have enabled the practical delivery of drugs such as insulin to blood circulation (see, e.g.,

Patton, J.S. et al., U.S. Patent No. 5,997,848). Pulmonary delivery has the advantage of administration without needle injection, and obviates the need for potentially toxic penetration enhancers.

Pharmaceutical compositions suitable for use in the invention include compositions wherein the active ingredients are contained in an effective amount to achieve the intended purpose. The determination of an effective dose is well within the capability of those skilled in the art.

Specialized forms of pharmaceutical compositions may be prepared for direct intracellular delivery of macromolecules comprising GBAP or fragments thereof. For example, liposome preparations containing a cell-impermeable macromolecule may promote cell fusion and intracellular delivery of the macromolecule. Alternatively, GBAP or a fragment thereof may be joined to a short cationic N-terminal portion from the HIV Tat-1 protein. Fusion proteins thus generated have been found to transduce into the cells of all tissues, including the brain, in a mouse model system (Schwarze, S.R. et al. (1999) Science 285:1569-1572).

For any compound, the therapeutically effective dose can be estimated initially either in cell culture assays, e.g., of neoplastic cells, or in animal models such as mice, rats, rabbits, dogs, monkeys, or pigs. An animal model may also be used to determine the appropriate concentration range and route of administration. Such information can then be used to determine useful doses and routes for administration in humans.

A therapeutically effective dose refers to that amount of active ingredient, for example GBAP or fragments thereof, antibodies of GBAP, and agonists, antagonists or inhibitors of GBAP, which ameliorates the symptoms or condition. Therapeutic efficacy and toxicity may be determined by standard pharmaceutical procedures in cell cultures or with experimental animals, such as by calculating the ED_{50} (the dose therapeutically effective in 50% of the population) or LD_{50} (the dose lethal to 50% of the population) statistics. The dose ratio of toxic to therapeutic effects is the therapeutic index, which can be expressed as the LD_{50}/ED_{50} ratio. Pharmaceutical compositions which exhibit large therapeutic indices are preferred. The data obtained from cell culture assays and animal studies are used to formulate a range of dosage for human use. The dosage contained in such compositions is preferably within a range of circulating concentrations that includes the ED_{50} with little or no toxicity. The dosage varies within this range depending upon the dosage form employed, the sensitivity of the patient, and the route of administration.

The exact dosage will be determined by the practitioner, in light of factors related to the subject requiring treatment. Dosage and administration are adjusted to provide sufficient levels of the active moiety or to maintain the desired effect. Factors which may be taken into account include the severity of the disease state, the general health of the subject, the age, weight, and gender of the subject, time and frequency of administration, drug combination(s), reaction sensitivities, and response to therapy. Long-acting pharmaceutical compositions may be administered every 3 to 4 days, every week, or

biweekly depending on the half-life and clearance rate of the particular formulation.

Normal dosage amounts may vary from about 0.1 μg to 100,000 μg, up to a total dose of about 1 gram, depending upon the route of administration. Guidance as to particular dosages and methods of delivery is provided in the literature and generally available to practitioners in the art. Those skilled in the art will employ different formulations for nucleotides than for proteins or their inhibitors. Similarly, delivery of polynucleotides or polypeptides will be specific to particular cells, conditions, locations, etc.

DIAGNOSTICS

In another embodiment, antibodies which specifically bind GBAP may be used for the diagnosis of disorders characterized by expression of GBAP, or in assays to monitor patients being treated with GBAP or agonists, antagonists, or inhibitors of GBAP. Antibodies useful for diagnostic purposes may be prepared in the same manner as described above for therapeutics. Diagnostic assays for GBAP include methods which utilize the antibody and a label to detect GBAP in human body fluids or in extracts of cells or tissues. The antibodies may be used with or without modification, and may be labeled by covalent or non-covalent attachment of a reporter molecule. A wide variety of reporter molecules, several of which are described above, are known in the art and may be used.

A variety of protocols for measuring GBAP, including ELISAs, RIAs, and FACS, are known in the art and provide a basis for diagnosing altered or abnormal levels of GBAP expression. Normal or standard values for GBAP expression are established by combining body fluids or cell extracts taken from normal mammalian subjects, for example, human subjects, with antibody to GBAP under conditions suitable for complex formation. The amount of standard complex formation may be quantitated by various methods, such as photometric means. Quantities of GBAP expressed in subject, control, and disease samples from biopsied tissues are compared with the standard values. Deviation between standard and subject values establishes the parameters for diagnosing disease.

In another embodiment of the invention, the polynucleotides encoding GBAP may be used for diagnostic purposes. The polynucleotides which may be used include oligonucleotide sequences, complementary RNA and DNA molecules, and PNAs. The polynucleotides may be used to detect and quantify gene expression in biopsied tissues in which expression of GBAP may be correlated with disease. The diagnostic assay may be used to determine absence, presence, and excess expression of GBAP, and to monitor regulation of GBAP levels during therapeutic intervention.

In one aspect, hybridization with PCR probes which are capable of detecting polynucleotide sequences, including genomic sequences, encoding GBAP or closely related molecules may be used to identify nucleic acid sequences which encode GBAP. The specificity of the probe, whether it is made from a highly specific region, e.g., the 5' regulatory region, or from a less specific region, e.g., a conserved motif, and the stringency of the hybridization or amplification will determine whether the

probe identifies only naturally occurring sequences encoding GBAP, allelic variants, or related sequences.

Probes may also be used for the detection of related sequences, and may have at least 50% sequence identity to any of the GBAP encoding sequences. The hybridization probes of the subject invention may be DNA or RNA and may be derived from the sequence of SEQ ID NO:67-132 or from genomic sequences including promoters, enhancers, and introns of the GBAP gene.

Means for producing specific hybridization probes for DNAs encoding GBAP include the cloning of polynucleotide sequences encoding GBAP or GBAP derivatives into vectors for the production of mRNA probes. Such vectors are known in the art, are commercially available, and may be used to synthesize RNA probes in vitro by means of the addition of the appropriate RNA polymerases and the appropriate labeled nucleotides. Hybridization probes may be labeled by a variety of reporter groups, for example, by radionuclides such as ^{32}P or ^{35}S, or by enzymatic labels, such as alkaline phosphatase coupled to the probe via avidin/biotin coupling systems, and the like.

Polynucleotide sequences encoding GBAP may be used for the diagnosis of disorders associated with expression of GBAP. Examples of such disorders include, but are not limited to, an immune system disorder such as inflammation, actinic keratosis, acquired immunodeficiency syndrome (AIDS), Addison's disease, adult respiratory distress syndrome, allergies, ankylosing spondylitis, amyloidosis, anemia, arteriosclerosis, asthma, atherosclerosis, autoimmune hemolytic anemia, autoimmune thyroiditis, bronchitis, bursitis, cholecystitis, cirrhosis, contact dermatitis, Crohn's disease, atopic dermatitis, dermatomyositis, diabetes mellitus, emphysema, erythroblastosis fetalis, erythema nodosum, atrophic gastritis, glomerulonephritis, Goodpasture's syndrome, gout, Graves' disease, Hashimoto's thyroiditis, paroxysmal nocturnal hemoglobinuria, hepatitis, hypereosinophilia, irritable bowel syndrome, episodic lymphopenia with lymphocytotoxins, mixed connective tissue disease (MCTD), multiple sclerosis, myasthenia gravis, myocardial or pericardial inflammation, myelofibrosis, osteoarthritis, osteoporosis, pancreatitis, polycythemia vera, polymyositis, psoriasis, Reiter's syndrome, rheumatoid arthritis, scleroderma, Sjögren's syndrome, systemic anaphylaxis, systemic lupus erythematosus, systemic sclerosis, primary thrombocythemia, thrombocytopenic purpura, ulcerative colitis, uveitis, Werner syndrome, complications of cancer, hemodialysis, and extracorporeal circulation, trauma, and hematopoietic cancer including lymphoma, leukemia, and myeloma; a reproductive disorder such as a disorder of prolactin production, infertility, including tubal disease, ovulatory defects, and endometriosis, a disruption of the estrous cycle, a disruption of the menstrual cycle, polycystic ovary syndrome, ovarian hyperstimulation syndrome, an endometrial or ovarian tumor, a uterine fibroid, autoimmune disorders, an ectopic pregnancy, and teratogenesis, cancer of the breast, fibrocystic breast disease, and galactorrhea, a disruption of spermatogenesis, abnormal sperm physiology, cancer of the testis, cancer of the prostate, benign

prostatic hyperplasia, prostatitis, Peyronie's disease, impotence, carcinoma of the male breast, and gynecomastia; a nervous system disorder such as epilepsy, ischemic cerebrovascular disease, stroke, cerebral neoplasms, Alzheimer's disease, Pick's disease, Huntington's disease, dementia, Parkinson's disease and other extrapyramidal disorders, amyotrophic lateral sclerosis and other motor neuron disorders, progressive neural muscular atrophy, retinitis pigmentosa, hereditary ataxias, multiple sclerosis and other demyelinating diseases, bacterial and viral meningitis, brain abscess, subdural empyema, epidural abscess, suppurative intracranial thrombophlebitis, myelitis and radiculitis, viral central nervous system disease, prion diseases including kuru, Creutzfeldt-Jakob disease, and Gerstmann-Straussler-Scheinker syndrome, fatal familial insomnia, nutritional and metabolic diseases of the nervous system, neurofibromatosis, tuberous sclerosis, cerebelloretinal hemangioblastomatosis, encephalotrigeminal syndrome, mental retardation and other developmental disorders of the central nervous system, cerebral palsy, neuroskeletal disorders, autonomic nervous system disorders, cranial nerve disorders, spinal cord diseases, muscular dystrophy and other neuromuscular disorders, peripheral nervous system disorders, dermatomyositis and polymyositis, inherited, metabolic, endocrine, and toxic myopathies, myasthenia gravis, periodic paralysis, mental disorders including mood, anxiety, and schizophrenic disorders, akathesia, amnesia, catatonia, diabetic neuropathy, tardive dyskinesia, dystonias, paranoid psychoses, postherpetic neuralgia, and Tourette's disorder; a cell signaling disorder including endocrine disorders such as disorders of the hypothalamus and pituitary resulting from lesions such as primary brain tumors, adenomas, infarction associated with pregnancy, hypophysectomy, aneurysms, vascular malformations, thrombosis, infections, immunological disorders, and complications due to head trauma; disorders associated with hyperpituitarism including acromegaly, giantism, and syndrome of inappropriate antidiuretic hormone (ADH) secretion (SIADH) often caused by benign adenoma; disorders associated with hypothyroidism including goiter, myxedema, acute thyroiditis associated with bacterial infection; disorders associated with hyperparathyroidism including Conn disease (chronic hypercalemia); pancreatic disorders such as Type I or Type II diabetes mellitus and associated complications; disorders associated with the adrenals such as hyperplasia, carcinoma, or adenoma of the adrenal cortex, hypertension associated with alkalosis; disorders associated with gonadal steroid hormones such as: in women, abnormal prolactin production, infertility, endometriosis, perturbations of the menstrual cycle, polycystic ovarian disease, hyperprolactinemia, isolated gonadotropin deficiency, amenorrhea, galactorrhea, hermaphroditism, hirsutism and virilization, breast cancer, and, in post-menopausal women, osteoporosis; and, in men, Leydig cell deficiency, male climacteric phase, and germinal cell aplasia, hypergonadal disorders associated with Leydig cell tumors, androgen resistance associated with absence of androgen receptors, syndrome of 5 α-reductase, and gynecomastia; and a cell proliferative disorder such as actinic keratosis, arteriosclerosis, atherosclerosis, bursitis, cirrhosis, hepatitis, mixed connective tissue disease (MCTD), myelofibrosis, paroxysmal nocturnal

hemoglobinuria, polycythemia vera, psoriasis, primary thrombocythemia, and cancers including adenocarcinoma, leukemia, lymphoma, melanoma, myeloma, sarcoma, teratocarcinoma, and, in particular, cancers of the adrenal gland, bladder, bone, bone marrow, brain, breast, cervix, gall bladder, ganglia, gastrointestinal tract, heart, kidney, liver, lung, muscle, ovary, pancreas, parathyroid, penis, prostate, salivary glands, skin, spleen, testis, thymus, thyroid, and uterus. The polynucleotide sequences encoding GBAP may be used in Southern or northern analysis, dot blot, or other membrane-based technologies; in PCR technologies; in dipstick, pin, and multiformat ELISA-like assays; and in microarrays utilizing fluids or tissues from patients to detect altered GBAP expression. Such qualitative or quantitative methods are well known in the art.

In a particular aspect, the nucleotide sequences encoding GBAP may be useful in assays that detect the presence of associated disorders, particularly those mentioned above. The nucleotide sequences encoding GBAP may be labeled by standard methods and added to a fluid or tissue sample from a patient under conditions suitable for the formation of hybridization complexes. After a suitable incubation period, the sample is washed and the signal is quantified and compared with a standard value. If the amount of signal in the patient sample is significantly altered in comparison to a control sample then the presence of altered levels of nucleotide sequences encoding GBAP in the sample indicates the presence of the associated disorder. Such assays may also be used to evaluate the efficacy of a particular therapeutic treatment regimen in animal studies, in clinical trials, or to monitor the treatment of an individual patient.

In order to provide a basis for the diagnosis of a disorder associated with expression of GBAP, a normal or standard profile for expression is established. This may be accomplished by combining body fluids or cell extracts taken from normal subjects, either animal or human, with a sequence, or a fragment thereof, encoding GBAP, under conditions suitable for hybridization or amplification. Standard hybridization may be quantified by comparing the values obtained from normal subjects with values from an experiment in which a known amount of a substantially purified polynucleotide is used. Standard values obtained in this manner may be compared with values obtained from samples from patients who are symptomatic for a disorder. Deviation from standard values is used to establish the presence of a disorder.

Once the presence of a disorder is established and a treatment protocol is initiated, hybridization assays may be repeated on a regular basis to determine if the level of expression in the patient begins to approximate that which is observed in the normal subject. The results obtained from successive assays may be used to show the efficacy of treatment over a period ranging from several days to months.

With respect to cancer, the presence of an abnormal amount of transcript (either under- or overexpressed) in biopsied tissue from an individual may indicate a predisposition for the development

of the disease, or may provide a means for detecting the disease prior to the appearance of actual clinical symptoms. A more definitive diagnosis of this type may allow health professionals to employ preventative measures or aggressive treatment earlier thereby preventing the development or further progression of the cancer.

Additional diagnostic uses for oligonucleotides designed from the sequences encoding GBAP may involve the use of PCR. These oligomers may be chemically synthesized, generated enzymatically, or produced in vitro. Oligomers will preferably contain a fragment of a polynucleotide encoding GBAP, or a fragment of a polynucleotide complementary to the polynucleotide encoding GBAP, and will be employed under optimized conditions for identification of a specific gene or condition. Oligomers may also be employed under less stringent conditions for detection or quantification of closely related DNA or RNA sequences.

In a particular aspect, oligonucleotide primers derived from the polynucleotide sequences encoding GBAP may be used to detect single nucleotide polymorphisms (SNPs). SNPs are substitutions, insertions and deletions that are a frequent cause of inherited or acquired genetic disease in humans. Methods of SNP detection include, but are not limited to, single-stranded conformation polymorphism (SSCP) and fluorescent SSCP (fSSCP) methods. In SSCP, oligonucleotide primers derived from the polynucleotide sequences encoding GBAP are used to amplify DNA using the polymerase chain reaction (PCR). The DNA may be derived, for example, from diseased or normal tissue, biopsy samples, bodily fluids, and the like. SNPs in the DNA cause differences in the secondary and tertiary structures of PCR products in single-stranded form, and these differences are detectable using gel electrophoresis in non-denaturing gels. In fSCCP, the oligonucleotide primers are fluorescently labeled, which allows detection of the amplimers in high-throughput equipment such as DNA sequencing machines. Additionally, sequence database analysis methods, termed in silico SNP (isSNP), are capable of identifying polymorphisms by comparing the sequence of individual overlapping DNA fragments which assemble into a common consensus sequence. These computer-based methods filter out sequence variations due to laboratory preparation of DNA and sequencing errors using statistical models and automated analyses of DNA sequence chromatograms. In the alternative, SNPs may be detected and characterized by mass spectrometry using, for example, the high throughput MASSARRAY system (Sequenom, Inc., San Diego CA).

Methods which may also be used to quantify the expression of GBAP include radiolabeling or biotinylating nucleotides, coamplification of a control nucleic acid, and interpolating results from standard curves. (See, e.g., Melby, P.C. et al. (1993) J. Immunol. Methods 159:235-244; Duplaa, C. et al. (1993) Anal. Biochem. 212:229-236.) The speed of quantitation of multiple samples may be accelerated by running the assay in a high-throughput format where the oligomer or polynucleotide of interest is presented in various dilutions and a spectrophotometric or colorimetric response gives rapid

quantitation.

In further embodiments, oligonucleotides or longer fragments derived from any of the polynucleotide sequences described herein may be used as elements on a microarray. The microarray can be used in transcript imaging techniques which monitor the relative expression levels of large numbers of genes simultaneously as described in Seilhamer, J.J. et al., "Comparative Gene Transcript Analysis," U.S. Patent No. 5,840,484, incorporated herein by reference. The microarray may also be used to identify genetic variants, mutations, and polymorphisms. This information may be used to determine gene function, to understand the genetic basis of a disorder, to diagnose a disorder, to monitor progression/regression of disease as a function of gene expression, and to develop and monitor the activities of therapeutic agents in the treatment of disease. In particular, this information may be used to develop a pharmacogenomic profile of a patient in order to select the most appropriate and effective treatment regimen for that patient. For example, therapeutic agents which are highly effective and display the fewest side effects may be selected for a patient based on his/her pharmacogenomic profile.

In another embodiment, antibodies specific for GBAP, or GBAP or fragments thereof may be used as elements on a microarray. The microarray may be used to monitor or measure protein-protein interactions, drug-target interactions, and gene expression profiles, as described above.

A particular embodiment relates to the use of the polynucleotides of the present invention to generate a transcript image of a tissue or cell type. A transcript image represents the global pattern of gene expression by a particular tissue or cell type. Global gene expression patterns are analyzed by quantifying the number of expressed genes and their relative abundance under given conditions and at a given time. (See Seilhamer et al., "Comparative Gene Transcript Analysis," U.S. Patent Number 5,840,484, expressly incorporated by reference herein.) Thus a transcript image may be generated by hybridizing the polynucleotides of the present invention or their complements to the totality of transcripts or reverse transcripts of a particular tissue or cell type. In one embodiment, the hybridization takes place in high-throughput format, wherein the polynucleotides of the present invention or their complements comprise a subset of a plurality of elements on a microarray. The resultant transcript image would provide a profile of gene activity.

Transcript images may be generated using transcripts isolated from tissues, cell lines, biopsies, or other biological samples. The transcript image may thus reflect gene expression in vivo, as in the case of a tissue or biopsy sample, or in vitro, as in the case of a cell line.

Transcript images which profile the expression of the polynucleotides of the present invention may also be used in conjunction with in vitro model systems and preclinical evaluation of pharmaceuticals, as well as toxicological testing of industrial and naturally-occurring environmental compounds. All compounds induce characteristic gene expression patterns, frequently termed molecular fingerprints or toxicant signatures, which are indicative of mechanisms of action and toxicity

(Nuwaysir, E.F. et al. (1999) Mol. Carcinog. 24:153-159; Steiner, S. and N.L. Anderson (2000) Toxicol. Lett. 112-113:467-471, expressly incorporated by reference herein). If a test compound has a signature similar to that of a compound with known toxicity, it is likely to share those toxic properties. These fingerprints or signatures are most useful and refined when they contain expression information from a large number of genes and gene families. Ideally, a genome-wide measurement of expression provides the highest quality signature. Even genes whose expression is not altered by any tested compounds are important as well, as the levels of expression of these genes are used to normalize the rest of the expression data. The normalization procedure is useful for comparison of expression data after treatment with different compounds. While the assignment of gene function to elements of a toxicant signature aids in interpretation of toxicity mechanisms, knowledge of gene function is not necessary for the statistical matching of signatures which leads to prediction of toxicity. (See, for example, Press Release 00-02 from the National Institute of Environmental Health Sciences, released February 29, 2000, available at http://www.niehs.nih.gov/oc/news/toxchip.htm.) Therefore, it is important and desirable in toxicological screening using toxicant signatures to include all expressed gene sequences.

In one embodiment, the toxicity of a test compound is assessed by treating a biological sample containing nucleic acids with the test compound. Nucleic acids that are expressed in the treated biological sample are hybridized with one or more probes specific to the polynucleotides of the present invention, so that transcript levels corresponding to the polynucleotides of the present invention may be quantified. The transcript levels in the treated biological sample are compared with levels in an untreated biological sample. Differences in the transcript levels between the two samples are indicative of a toxic response caused by the test compound in the treated sample.

Another particular embodiment relates to the use of the polypeptide sequences of the present invention to analyze the proteome of a tissue or cell type. The term proteome refers to the global pattern of protein expression in a particular tissue or cell type. Each protein component of a proteome can be subjected individually to further analysis. Proteome expression patterns, or profiles, are analyzed by quantifying the number of expressed proteins and their relative abundance under given conditions and at a given time. A profile of a cell's proteome may thus be generated by separating and analyzing the polypeptides of a particular tissue or cell type. In one embodiment, the separation is achieved using two-dimensional gel electrophoresis, in which proteins from a sample are separated by isoelectric focusing in the first dimension, and then according to molecular weight by sodium dodecyl sulfate slab gel electrophoresis in the second dimension (Steiner and Anderson, supra). The proteins are visualized in the gel as discrete and uniquely positioned spots, typically by staining the gel with an agent such as Coomassie Blue or silver or fluorescent stains. The optical density of each protein spot is generally proportional to the level of the protein in the sample. The optical densities of equivalently

positioned protein spots from different samples, for example, from biological samples either treated or untreated with a test compound or therapeutic agent, are compared to identify any changes in protein spot density related to the treatment. The proteins in the spots are partially sequenced using, for example, standard methods employing chemical or enzymatic cleavage followed by mass spectrometry. The identity of the protein in a spot may be determined by comparing its partial sequence, preferably of at least 5 contiguous amino acid residues, to the polypeptide sequences of the present invention. In some cases, further sequence data may be obtained for definitive protein identification.

A proteomic profile may also be generated using antibodies specific for GBAP to quantify the levels of GBAP expression. In one embodiment, the antibodies are used as elements on a microarray, and protein expression levels are quantified by exposing the microarray to the sample and detecting the levels of protein bound to each array element (Lueking, A. et al. (1999) Anal. Biochem. 270:103-111; Mendoze, L.G. et al. (1999) Biotechniques 27:778-788). Detection may be performed by a variety of methods known in the art, for example, by reacting the proteins in the sample with a thiol- or amino-reactive fluorescent compound and detecting the amount of fluorescence bound at each array element.

Toxicant signatures at the proteome level are also useful for toxicological screening, and should be analyzed in parallel with toxicant signatures at the transcript level. There is a poor correlation between transcript and protein abundances for some proteins in some tissues (Anderson, N.L. and J. Seilhamer (1997) Electrophoresis 18:533-537), so proteome toxicant signatures may be useful in the analysis of compounds which do not significantly affect the transcript image, but which alter the proteomic profile. In addition, the analysis of transcripts in body fluids is difficult, due to rapid degradation of mRNA, so proteomic profiling may be more reliable and informative in such cases.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins that are expressed in the treated biological sample are separated so that the amount of each protein can be quantified. The amount of each protein is compared to the amount of the corresponding protein in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample. Individual proteins are identified by sequencing the amino acid residues of the individual proteins and comparing these partial sequences to the polypeptides of the present invention.

In another embodiment, the toxicity of a test compound is assessed by treating a biological sample containing proteins with the test compound. Proteins from the biological sample are incubated with antibodies specific to the polypeptides of the present invention. The amount of protein recognized by the antibodies is quantified. The amount of protein in the treated biological sample is compared with the amount in an untreated biological sample. A difference in the amount of protein between the two samples is indicative of a toxic response to the test compound in the treated sample.

Microarrays may be prepared, used, and analyzed using methods known in the art. (See, e.g.,

Brennan, T.M. et al. (1995) U.S. Patent No. 5,474,796; Schena, M. et al. (1996) Proc. Natl. Acad. Sci. USA 93:10614-10619; Baldeschweiler et al. (1995) PCT application WO95/251116; Shalon, D. et al. (1995) PCT application WO95/35505; Heller, R.A. et al. (1997) Proc. Natl. Acad. Sci. USA 94:2150-2155; and Heller, M.J. et al. (1997) U.S. Patent No. 5,605,662.) Various types of microarrays are well known and thoroughly described in DNA Microarrays: A Practical Approach, M. Schena, ed. (1999) Oxford University Press, London, hereby expressly incorporated by reference.

In another embodiment of the invention, nucleic acid sequences encoding GBAP may be used to generate hybridization probes useful in mapping the naturally occurring genomic sequence. Either coding or noncoding sequences may be used, and in some instances, noncoding sequences may be preferable over coding sequences. For example, conservation of a coding sequence among members of a multi-gene family may potentially cause undesired cross hybridization during chromosomal mapping. The sequences may be mapped to a particular chromosome, to a specific region of a chromosome, or to artificial chromosome constructions, e.g., human artificial chromosomes (HACs), yeast artificial chromosomes (YACs), bacterial artificial chromosomes (BACs), bacterial P1 constructions, or single chromosome cDNA libraries. (See, e.g., Harrington, J.J. et al. (1997) Nat. Genet. 15:345-355; Price, C.M. (1993) Blood Rev. 7:127-134; and Trask, B.J. (1991) Trends Genet. 7:149-154.) Once mapped, the nucleic acid sequences of the invention may be used to develop genetic linkage maps, for example, which correlate the inheritance of a disease state with the inheritance of a particular chromosome region or restriction fragment length polymorphism (RFLP). (See, e.g., Lander, E.S. and D. Botstein (1986) Proc. Natl. Acad. Sci. USA 83:7353-7357.)

Fluorescent in situ hybridization (FISH) may be correlated with other physical and genetic map data. (See, e.g., Heinz-Ulrich, et al. (1995) in Meyers, supra, pp. 965-968.) Examples of genetic map data can be found in various scientific journals or at the Online Mendelian Inheritance in Man (OMIM) World Wide Web site. Correlation between the location of the gene encoding GBAP on a physical map and a specific disorder, or a predisposition to a specific disorder, may help define the region of DNA associated with that disorder and thus may further positional cloning efforts.

In situ hybridization of chromosomal preparations and physical mapping techniques, such as linkage analysis using established chromosomal markers, may be used for extending genetic maps. Often the placement of a gene on the chromosome of another mammalian species, such as mouse, may reveal associated markers even if the exact chromosomal locus is not known. This information is valuable to investigators searching for disease genes using positional cloning or other gene discovery techniques. Once the gene or genes responsible for a disease or syndrome have been crudely localized by genetic linkage to a particular genomic region, e.g., ataxia-telangiectasia to 11q22-23, any sequences mapping to that area may represent associated or regulatory genes for further investigation. (See, e.g., Gatti, R.A. et al. (1988) Nature 336:577-580.) The nucleotide sequence of the instant invention may

also be used to detect differences in the chromosomal location due to translocation, inversion, etc., among normal, carrier, or affected individuals.

In another embodiment of the invention, GBAP, its catalytic or immunogenic fragments, or oligopeptides thereof can be used for screening libraries of compounds in any of a variety of drug screening techniques. The fragment employed in such screening may be free in solution, affixed to a solid support, borne on a cell surface, or located intracellularly. The formation of binding complexes between GBAP and the agent being tested may be measured.

Another technique for drug screening provides for high throughput screening of compounds having suitable binding affinity to the protein of interest. (See, e.g., Geysen, et al. (1984) PCT application WO84/03564.) In this method, large numbers of different small test compounds are synthesized on a solid substrate. The test compounds are reacted with GBAP, or fragments thereof, and washed. Bound GBAP is then detected by methods well known in the art. Purified GBAP can also be coated directly onto plates for use in the aforementioned drug screening techniques. Alternatively, non-neutralizing antibodies can be used to capture the peptide and immobilize it on a solid support.

In another embodiment, one may use competitive drug screening assays in which neutralizing antibodies capable of binding GBAP specifically compete with a test compound for binding GBAP. In this manner, antibodies can be used to detect the presence of any peptide which shares one or more antigenic determinants with GBAP.

In additional embodiments, the nucleotide sequences which encode GBAP may be used in any molecular biology techniques that have yet to be developed, provided the new techniques rely on properties of nucleotide sequences that are currently known, including, but not limited to, such properties as the triplet genetic code and specific base pair interactions.

Without further elaboration, it is believed that one skilled in the art can, using the preceding description, utilize the present invention to its fullest extent. The following preferred specific embodiments are, therefore, to be construed as merely illustrative, and not limitative of the remainder of the disclosure in any way whatsoever.

The disclosures of all patents, applications and publications, mentioned above and below, in particular U.S. Ser. No. 60/144,595, U.S Ser. No. 60/150,460, and U.S. Ser. No. 60/159,849, are hereby expressly incorporated by reference.

EXAMPLES

I. Construction of cDNA Libraries

RNA was purchased from Clontech or isolated from tissues described in Table 4. Some tissues were homogenized and lysed in guanidinium isothiocyanate, while others were homogenized and lysed in phenol or in a suitable mixture of denaturants, such as TRIZOL (Life Technologies), a monophasic

solution of phenol and guanidine isothiocyanate. The resulting lysates were centrifuged over CsCl cushions or extracted with chloroform. RNA was precipitated from the lysates with either isopropanol or sodium acetate and ethanol, or by other routine methods.

Phenol extraction and precipitation of RNA were repeated as necessary to increase RNA purity. In some cases, RNA was treated with DNase. For most libraries, poly(A+) RNA was isolated using oligo d(T)-coupled paramagnetic particles (Promega), OLIGOTEX latex particles (QIAGEN, Chatsworth CA), or an OLIGOTEX mRNA purification kit (QIAGEN). Alternatively, RNA was isolated directly from tissue lysates using other RNA isolation kits, e.g., the POLY(A)PURE mRNA purification kit (Ambion, Austin TX).

In some cases, Stratagene was provided with RNA and constructed the corresponding cDNA libraries. Otherwise, cDNA was synthesized and cDNA libraries were constructed with the UNIZAP vector system (Stratagene) or SUPERSCRIPT plasmid system (Life Technologies), using the recommended procedures or similar methods known in the art. (See, e.g., Ausubel, 1997, supra, units 5.1-6.6.) Reverse transcription was initiated using oligo d(T) or random primers. Synthetic oligonucleotide adapters were ligated to double stranded cDNA, and the cDNA was digested with the appropriate restriction enzyme or enzymes. For most libraries, the cDNA was size-selected (300-1000 bp) using SEPHACRYL S1000, SEPHAROSE CL2B, or SEPHAROSE CL4B column chromatography (Amersham Pharmacia Biotech) or preparative agarose gel electrophoresis. cDNAs were ligated into compatible restriction enzyme sites of the polylinker of a suitable plasmid, e.g., PBLUESCRIPT plasmid (Stratagene), PSPORT1 plasmid (Life Technologies), pcDNA2.1 plasmid (Invitrogen, Carlsbad CA), or pINCY plasmid (Incyte Genomics, Palo Alto CA). Recombinant plasmids were transformed into competent E. coli cells including XL1-Blue, XL1-BlueMRF, or SOLR from Stratagene or DH5α, DH10B, or ElectroMAX DH10B from Life Technologies.

II. Isolation of cDNA Clones

Plasmids obtained as described in Example I were recovered from host cells by in vivo excision using the UNIZAP vector system (Stratagene) or by cell lysis. Plasmids were purified using at least one of the following: a Magic or WIZARD Minipreps DNA purification system (Promega); an AGTC Miniprep purification kit (Edge Biosystems, Gaithersburg MD); and QIAWELL 8 Plasmid, QIAWELL 8 Plus Plasmid, QIAWELL 8 Ultra Plasmid purification systems or the R.E.A.L. PREP 96 plasmid purification kit from QIAGEN. Following precipitation, plasmids were resuspended in 0.1 ml of distilled water and stored, with or without lyophilization, at 4°C.

Alternatively, plasmid DNA was amplified from host cell lysates using direct link PCR in a high-throughput format (Rao, V.B. (1994) Anal. Biochem. 216:1-14). Host cell lysis and thermal cycling steps were carried out in a single reaction mixture. Samples were processed and stored in 384-well plates, and the concentration of amplified plasmid DNA was quantified fluorometrically using

PICOGREEN dye (Molecular Probes, Eugene OR) and a FLUOROSKAN II fluorescence scanner (Labsystems Oy, Helsinki, Finland).

III. Sequencing and Analysis

Incyte cDNA recovered in plasmids as described in Example II were sequenced as follows. Sequencing reactions were processed using standard methods or high-throughput instrumentation such as the ABI CATALYST 800 (PE Biosystems) thermal cycler or the PTC-200 thermal cycler (MJ Research) in conjunction with the HYDRA microdispenser (Robbins Scientific) or the MICROLAB 2200 (Hamilton) liquid transfer system. cDNA sequencing reactions were prepared using reagents provided by Amersham Pharmacia Biotech or supplied in ABI sequencing kits such as the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems). Electrophoretic separation of cDNA sequencing reactions and detection of labeled polynucleotides were carried out using the MEGABACE 1000 DNA sequencing system (Molecular Dynamics); the ABI PRISM 373 or 377 sequencing system (PE Biosystems) in conjunction with standard ABI protocols and base calling software; or other sequence analysis systems known in the art. Reading frames within the cDNA sequences were identified using standard methods (reviewed in Ausubel, 1997, supra, unit 7.7). Some of the cDNA sequences were selected for extension using the techniques disclosed in Example VI.

The polynucleotide sequences derived from cDNA sequencing were assembled and analyzed using a combination of software programs which utilize algorithms well known to those skilled in the art. Table 5 summarizes the tools, programs, and algorithms used and provides applicable descriptions, references, and threshold parameters. The first column of Table 5 shows the tools, programs, and algorithms used, the second column provides brief descriptions thereof, the third column presents appropriate references, all of which are incorporated by reference herein in their entirety, and the fourth column presents, where applicable, the scores, probability values, and other parameters used to evaluate the strength of a match between two sequences (the higher the score, the greater the homology between two sequences). Sequences were analyzed using MACDNASIS PRO software (Hitachi Software Engineering, South San Francisco CA) and LASERGENE software (DNASTAR). Polynucleotide and polypeptide sequence alignments were generated using the default parameters specified by the clustal algorithm as incorporated into the MEGALIGN multisequence alignment program (DNASTAR), which also calculates the percent identity between aligned sequences.

The polynucleotide sequences were validated by removing vector, linker, and polyA sequences and by masking ambiguous bases, using algorithms and programs based on BLAST, dynamic programing, and dinucleotide nearest neighbor analysis. The sequences were then queried against a selection of public databases such as the GenBank primate, rodent, mammalian, vertebrate, and eukaryote databases, and BLOCKS, PRINTS, DOMO, PRODOM, and PFAM to acquire annotation using programs based on BLAST, FASTA, and BLIMPS. The sequences were assembled into full

length polynucleotide sequences using programs based on Phred, Phrap, and Consed, and were screened for open reading frames using programs based on GeneMark, BLAST, and FASTA. The full length polynucleotide sequences were translated to derive the corresponding full length amino acid sequences, and these full length sequences were subsequently analyzed by querying against databases such as the GenBank databases (described above), SwissProt, BLOCKS, PRINTS, DOMO, PRODOM, Prosite, and Hidden Markov Model (HMM)-based protein family databases such as PFAM. HMM is a probabilistic approach which analyzes consensus primary structures of gene families. (See, e.g., Eddy, S.R. (1996) Curr. Opin. Struct. Biol. 6:361-365.)

The programs described above for the assembly and analysis of full length polynucleotide and amino acid sequences were also used to identify polynucleotide sequence fragments from SEQ ID NO:67-132. Fragments from about 20 to about 4000 nucleotides which are useful in hybridization and amplification technologies were described in The Invention section above.

IV. Analysis of Polynucleotide Expression

Northern analysis is a laboratory technique used to detect the presence of a transcript of a gene and involves the hybridization of a labeled nucleotide sequence to a membrane on which RNAs from a particular cell type or tissue have been bound. (See, e.g., Sambrook, <u>supra</u>, ch. 7; Ausubel, 1995, <u>supra</u>, ch. 4 and 16.)

Analogous computer techniques applying BLAST were used to search for identical or related molecules in cDNA databases such as GenBank or LIFESEQ (Incyte Genomics). This analysis is much faster than multiple membrane-based hybridizations. In addition, the sensitivity of the computer search can be modified to determine whether any particular match is categorized as exact or similar. The basis of the search is the product score, which is defined as:

$$\frac{\text{BLAST Score x Percent Identity}}{5 \text{ x minimum } \{\text{length(Seq. 1), length(Seq. 2)}\}}$$

The product score takes into account both the degree of similarity between two sequences and the length of the sequence match. The product score is a normalized value between 0 and 100, and is calculated as follows: the BLAST score is multiplied by the percent nucleotide identity and the product is divided by (5 times the length of the shorter of the two sequences). The BLAST score is calculated by assigning a score of +5 for every base that matches in a high-scoring segment pair (HSP), and -4 for every mismatch. Two sequences may share more than one HSP (separated by gaps). If there is more than one HSP, then the pair with the highest BLAST score is used to calculate the product score. The product score represents a balance between fractional overlap and quality in a BLAST alignment. For example, a product score of 100 is produced only for 100% identity over the entire length of the shorter of the two sequences being compared. A product score of 70 is produced either by 100% identity and

70% overlap at one end, or by 88% identity and 100% overlap at the other. A product score of 50 is produced either by 100% identity and 50% overlap at one end, or 79% identity and 100% overlap.

The results of northern analyses are reported as a percentage distribution of libraries in which the transcript encoding GBAP occurred. Analysis involved the categorization of cDNA libraries by organ/tissue and disease. The organ/tissue categories included cardiovascular, dermatologic, developmental, endocrine, gastrointestinal, hematopoietic/immune, musculoskeletal, nervous, reproductive, and urologic. The disease/condition categories included cancer, inflammation, trauma, cell proliferation, neurological, and pooled. For each category, the number of libraries expressing the sequence of interest was counted and divided by the total number of libraries across all categories. Percentage values of tissue-specific and disease- or condition-specific expression are reported in Table 3.

V. Chromosomal Mapping of GBAP Encoding Polynucleotides

The cDNA sequences which were used to assemble SEQ ID NO:67-132 were compared with sequences from the Incyte LIFESEQ database and public domain databases using BLAST and other implementations of the Smith-Waterman algorithm. Sequences from these databases that matched SEQ ID NO:67-132 were assembled into clusters of contiguous and overlapping sequences using assembly algorithms such as Phrap (Table 5). Radiation hybrid and genetic mapping data available from public resources such as the Stanford Human Genome Center (SHGC), Whitehead Institute for Genome Research (WIGR), and Généthon were used to determine if any of the clustered sequences had been previously mapped. Inclusion of a mapped sequence in a cluster resulted in the assignment of all sequences of that cluster, including its particular SEQ ID NO:, to that map location.

The genetic map locations of SEQ ID NO:70, 74, 75, 77, 80, 86, 87, 90, 92, 93, 94, 97, 101, 106, 109, 111, 112, 113, 115, 117, 118, 121, and 128 are described in The Invention as ranges, or intervals, of human chromosomes. More than one map location is reported for SEQ ID NO:94, 101, 109, 111, and 115, indicating that previously mapped sequences having similarity, but not complete identity, to SEQ ID NO:94, 101, 109, 111, and 115 were assembled into their respective clusters. The map position of an interval, in centiMorgans, is measured relative to the terminus of the chromosome's p-arm. (The centiMorgan (cM) is a unit of measurement based on recombination frequencies between chromosomal markers. On average, 1 cM is roughly equivalent to 1 megabase (Mb) of DNA in humans, although this can vary widely due to hot and cold spots of recombination.) The cM distances are based on genetic markers mapped by Généthon which provide boundaries for radiation hybrid markers whose sequences were included in each of the clusters.

VI. Extension of GBAP Encoding Polynucleotides

The full length nucleic acid sequences of SEQ ID NO:67-132 were produced by extension of an appropriate fragment of the full length molecule using oligonucleotide primers designed from this

fragment. One primer was synthesized to initiate 5' extension of the known fragment, and the other primer, to initiate 3' extension of the known fragment. The initial primers were designed using OLIGO 4.06 software (National Biosciences), or another appropriate program, to be about 22 to 30 nucleotides in length, to have a GC content of about 50% or more, and to anneal to the target sequence at temperatures of about 68°C to about 72°C. Any stretch of nucleotides which would result in hairpin structures and primer-primer dimerizations was avoided.

Selected human cDNA libraries were used to extend the sequence. If more than one extension was necessary or desired, additional or nested sets of primers were designed.

High fidelity amplification was obtained by PCR using methods well known in the art. PCR was performed in 96-well plates using the PTC-200 thermal cycler (MJ Research, Inc.). The reaction mix contained DNA template, 200 nmol of each primer, reaction buffer containing Mg^{2+}, $(NH_4)_2SO_4$, and β-mercaptoethanol, Taq DNA polymerase (Amersham Pharmacia Biotech), ELONGASE enzyme (Life Technologies), and Pfu DNA polymerase (Stratagene), with the following parameters for primer pair PCI A and PCI B: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C. In the alternative, the parameters for primer pair T7 and SK+ were as follows: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 57°C, 1 min; Step 4: 68°C, 2 min; Step 5: Steps 2, 3, and 4 repeated 20 times; Step 6: 68°C, 5 min; Step 7: storage at 4°C.

The concentration of DNA in each well was determined by dispensing 100 µl PICOGREEN quantitation reagent (0.25% (v/v) PICOGREEN; Molecular Probes, Eugene OR) dissolved in 1X TE and 0.5 µl of undiluted PCR product into each well of an opaque fluorimeter plate (Corning Costar, Acton MA), allowing the DNA to bind to the reagent. The plate was scanned in a Fluoroskan II (Labsystems Oy, Helsinki, Finland) to measure the fluorescence of the sample and to quantify the concentration of DNA. A 5 µl to 10 µl aliquot of the reaction mixture was analyzed by electrophoresis on a 1 % agarose mini-gel to determine which reactions were successful in extending the sequence.

The extended nucleotides were desalted and concentrated, transferred to 384-well plates, digested with CviJI cholera virus endonuclease (Molecular Biology Research, Madison WI), and sonicated or sheared prior to religation into pUC 18 vector (Amersham Pharmacia Biotech). For shotgun sequencing, the digested nucleotides were separated on low concentration (0.6 to 0.8%) agarose gels, fragments were excised, and agar digested with Agar ACE (Promega). Extended clones were religated using T4 ligase (New England Biolabs, Beverly MA) into pUC 18 vector (Amersham Pharmacia Biotech), treated with Pfu DNA polymerase (Stratagene) to fill-in restriction site overhangs, and transfected into competent E. coli cells. Transformed cells were selected on antibiotic-containing media, and individual colonies were picked and cultured overnight at 37°C in 384-well plates in LB/2x carb liquid media.

The cells were lysed, and DNA was amplified by PCR using Taq DNA polymerase (Amersham Pharmacia Biotech) and Pfu DNA polymerase (Stratagene) with the following parameters: Step 1: 94°C, 3 min; Step 2: 94°C, 15 sec; Step 3: 60°C, 1 min; Step 4: 72°C, 2 min; Step 5: steps 2, 3, and 4 repeated 29 times; Step 6: 72°C, 5 min; Step 7: storage at 4°C. DNA was quantified by PICOGREEN reagent (Molecular Probes) as described above. Samples with low DNA recoveries were reamplified using the same conditions as described above. Samples were diluted with 20% dimethysulfoxide (1:2, v/v), and sequenced using DYENAMIC energy transfer sequencing primers and the DYENAMIC DIRECT kit (Amersham Pharmacia Biotech) or the ABI PRISM BIGDYE Terminator cycle sequencing ready reaction kit (PE Biosystems).

In like manner, the polynucleotide sequences of SEQ ID NO:67-132 are used to obtain 5' regulatory sequences using the procedure above, along with oligonucleotides designed for such extension, and an appropriate genomic library.

VII. Labeling and Use of Individual Hybridization Probes

Hybridization probes derived from SEQ ID NO:67-132 are employed to screen cDNAs, genomic DNAs, or mRNAs. Although the labeling of oligonucleotides, consisting of about 20 base pairs, is specifically described, essentially the same procedure is used with larger nucleotide fragments. Oligonucleotides are designed using state-of-the-art software such as OLIGO 4.06 software (National Biosciences) and labeled by combining 50 pmol of each oligomer, 250 μCi of [γ-^{32}P] adenosine triphosphate (Amersham Pharmacia Biotech), and T4 polynucleotide kinase (DuPont NEN, Boston MA). The labeled oligonucleotides are substantially purified using a SEPHADEX G-25 superfine size exclusion dextran bead column (Amersham Pharmacia Biotech). An aliquot containing 10^7 counts per minute of the labeled probe is used in a typical membrane-based hybridization analysis of human genomic DNA digested with one of the following endonucleases: Ase I, Bgl II, Eco RI, Pst I, Xba I, or Pvu II (DuPont NEN).

The DNA from each digest is fractionated on a 0.7% agarose gel and transferred to nylon membranes (Nytran Plus, Schleicher & Schuell, Durham NH). Hybridization is carried out for 16 hours at 40°C. To remove nonspecific signals, blots are sequentially washed at room temperature under conditions of up to, for example, 0.1 x saline sodium citrate and 0.5% sodium dodecyl sulfate. Hybridization patterns are visualized using autoradiography or an alternative imaging means and compared.

VIII. Microarrays

The linkage or synthesis of array elements upon a microarray can be achieved utilizing photolithography, piezoelectric printing (ink-jet printing, See, e.g., Baldeschweiler, supra), mechanical microspotting technologies, and derivatives thereof. The substrate in each of the aforementioned technologies should be uniform and solid with a non-porous surface (Schena (1999), supra). Suggested

substrates include silicon, silica, glass slides, glass chips, and silicon wafers. Alternatively, a procedure analogous to a dot or slot blot may also be used to arrange and link elements to the surface of a substrate using thermal, UV, chemical, or mechanical bonding procedures. A typical array may be produced using available methods and machines well known to those of ordinary skill in the art and may contain any appropriate number of elements. (See, e.g., Schena, M. et al. (1995) Science 270:467-470; Shalon, D. et al. (1996) Genome Res. 6:639-645; Marshall, A. and J. Hodgson (1998) Nat. Biotechnol. 16:27-31.)

Full length cDNAs, Expressed Sequence Tags (ESTs), or fragments or oligomers thereof may comprise the elements of the microarray. Fragments or oligomers suitable for hybridization can be selected using software well known in the art such as LASERGENE software (DNASTAR). The array elements are hybridized with polynucleotides in a biological sample. The polynucleotides in the biological sample are conjugated to a fluorescent label or other molecular tag for ease of detection. After hybridization, nonhybridized nucleotides from the biological sample are removed, and a fluorescence scanner is used to detect hybridization at each array element. Alternatively, laser desorbtion and mass spectrometry may be used for detection of hybridization. The degree of complementarity and the relative abundance of each polynucleotide which hybridizes to an element on the microarray may be assessed. In one embodiment, microarray preparation and usage is described in detail below.

Tissue or Cell Sample Preparation

Total RNA is isolated from tissue samples using the guanidinium thiocyanate method and poly(A)$^+$ RNA is purified using the oligo-(dT) cellulose method. Each poly(A)$^+$ RNA sample is reverse transcribed using MMLV reverse-transcriptase, 0.05 pg/µl oligo-(dT) primer (21mer), 1X first strand buffer, 0.03 units/µl RNase inhibitor, 500 µM dATP, 500 µM dGTP, 500 µM dTTP, 40 µM dCTP, 40 µM dCTP-Cy3 (BDS) or dCTP-Cy5 (Amersham Pharmacia Biotech). The reverse transcription reaction is performed in a 25 ml volume containing 200 ng poly(A)$^+$ RNA with GEMBRIGHT kits (Incyte). Specific control poly(A)$^+$ RNAs are synthesized by in vitro transcription from non-coding yeast genomic DNA. After incubation at 37°C for 2 hr, each reaction sample (one with Cy3 and another with Cy5 labeling) is treated with 2.5 ml of 0.5M sodium hydroxide and incubated for 20 minutes at 85°C to the stop the reaction and degrade the RNA. Samples are purified using two successive CHROMA SPIN 30 gel filtration spin columns (CLONTECH Laboratories, Inc. (CLONTECH), Palo Alto CA) and after combining, both reaction samples are ethanol precipitated using 1 ml of glycogen (1 mg/ml), 60 ml sodium acetate, and 300 ml of 100% ethanol. The sample is then dried to completion using a SpeedVAC (Savant Instruments Inc., Holbrook NY) and resuspended in 14 µl 5X SSC/0.2% SDS.

Microarray Preparation

Sequences of the present invention are used to generate array elements. Each array element is amplified from bacterial cells containing vectors with cloned cDNA inserts. PCR amplification uses primers complementary to the vector sequences flanking the cDNA insert. Array elements are amplified in thirty cycles of PCR from an initial quantity of 1-2 ng to a final quantity greater than 5 μg. Amplified array elements are then purified using SEPHACRYL-400 (Amersham Pharmacia Biotech).

Purified array elements are immobilized on polymer-coated glass slides. Glass microscope slides (Corning) are cleaned by ultrasound in 0.1% SDS and acetone, with extensive distilled water washes between and after treatments. Glass slides are etched in 4% hydrofluoric acid (VWR Scientific Products Corporation (VWR), West Chester PA), washed extensively in distilled water, and coated with 0.05% aminopropyl silane (Sigma) in 95% ethanol. Coated slides are cured in a 110°C oven.

Array elements are applied to the coated glass substrate using a procedure described in US Patent No. 5,807,522, incorporated herein by reference. 1 μl of the array element DNA, at an average concentration of 100 ng/μl, is loaded into the open capillary printing element by a high-speed robotic apparatus. The apparatus then deposits about 5 nl of array element sample per slide.

Microarrays are UV-crosslinked using a STRATALINKER UV-crosslinker (Stratagene). Microarrays are washed at room temperature once in 0.2% SDS and three times in distilled water. Non-specific binding sites are blocked by incubation of microarrays in 0.2% casein in phosphate buffered saline (PBS) (Tropix, Inc., Bedford MA) for 30 minutes at 60°C followed by washes in 0.2% SDS and distilled water as before.

Hybridization

Hybridization reactions contain 9 μl of sample mixture consisting of 0.2 μg each of Cy3 and Cy5 labeled cDNA synthesis products in 5X SSC, 0.2% SDS hybridization buffer. The sample mixture is heated to 65°C for 5 minutes and is aliquoted onto the microarray surface and covered with an 1.8 cm^2 coverslip. The arrays are transferred to a waterproof chamber having a cavity just slightly larger than a microscope slide. The chamber is kept at 100% humidity internally by the addition of 140 μl of 5X SSC in a corner of the chamber. The chamber containing the arrays is incubated for about 6.5 hours at 60°C. The arrays are washed for 10 min at 45°C in a first wash buffer (1X SSC, 0.1% SDS), three times for 10 minutes each at 45°C in a second wash buffer (0.1X SSC), and dried.

Detection

Reporter-labeled hybridization complexes are detected with a microscope equipped with an Innova 70 mixed gas 10 W laser (Coherent, Inc., Santa Clara CA) capable of generating spectral lines at 488 nm for excitation of Cy3 and at 632 nm for excitation of Cy5. The excitation laser light is focused on the array using a 20X microscope objective (Nikon, Inc., Melville NY). The slide

containing the array is placed on a computer-controlled X-Y stage on the microscope and raster-scanned past the objective. The 1.8 cm x 1.8 cm array used in the present example is scanned with a resolution of 20 micrometers.

In two separate scans, a mixed gas multiline laser excites the two fluorophores sequentially. Emitted light is split, based on wavelength, into two photomultiplier tube detectors (PMT R1477, Hamamatsu Photonics Systems, Bridgewater NJ) corresponding to the two fluorophores. Appropriate filters positioned between the array and the photomultiplier tubes are used to filter the signals. The emission maxima of the fluorophores used are 565 nm for Cy3 and 650 nm for Cy5. Each array is typically scanned twice, one scan per fluorophore using the appropriate filters at the laser source, although the apparatus is capable of recording the spectra from both fluorophores simultaneously.

The sensitivity of the scans is typically calibrated using the signal intensity generated by a cDNA control species added to the sample mixture at a known concentration. A specific location on the array contains a complementary DNA sequence, allowing the intensity of the signal at that location to be correlated with a weight ratio of hybridizing species of 1:100,000. When two samples from different sources (e.g., representing test and control cells), each labeled with a different fluorophore, are hybridized to a single array for the purpose of identifying genes that are differentially expressed, the calibration is done by labeling samples of the calibrating cDNA with the two fluorophores and adding identical amounts of each to the hybridization mixture.

The output of the photomultiplier tube is digitized using a 12-bit RTI-835H analog-to-digital (A/D) conversion board (Analog Devices, Inc., Norwood MA) installed in an IBM-compatible PC computer. The digitized data are displayed as an image where the signal intensity is mapped using a linear 20-color transformation to a pseudocolor scale ranging from blue (low signal) to red (high signal). The data is also analyzed quantitatively. Where two different fluorophores are excited and measured simultaneously, the data are first corrected for optical crosstalk (due to overlapping emission spectra) between the fluorophores using each fluorophore's emission spectrum.

A grid is superimposed over the fluorescence signal image such that the signal from each spot is centered in each element of the grid. The fluorescence signal within each element is then integrated to obtain a numerical value corresponding to the average intensity of the signal. The software used for signal analysis is the GEMTOOLS gene expression analysis program (Incyte).

IX. Complementary Polynucleotides

Sequences complementary to the GBAP-encoding sequences, or any parts thereof, are used to detect, decrease, or inhibit expression of naturally occurring GBAP. Although use of oligonucleotides comprising from about 15 to 30 base pairs is described, essentially the same procedure is used with smaller or with larger sequence fragments. Appropriate oligonucleotides are designed using OLIGO 4.06 software (National Biosciences) and the coding sequence of GBAP. To inhibit transcription, a

complementary oligonucleotide is designed from the most unique 5' sequence and used to prevent promoter binding to the coding sequence. To inhibit translation, a complementary oligonucleotide is designed to prevent ribosomal binding to the GBAP-encoding transcript.

X. Expression of GBAP

Expression and purification of GBAP is achieved using bacterial or virus-based expression systems. For expression of GBAP in bacteria, cDNA is subcloned into an appropriate vector containing an antibiotic resistance gene and an inducible promoter that directs high levels of cDNA transcription. Examples of such promoters include, but are not limited to, the *trp-lac* (*tac*) hybrid promoter and the T5 or T7 bacteriophage promoter in conjunction with the *lac* operator regulatory element. Recombinant vectors are transformed into suitable bacterial hosts, e.g., BL21(DE3). Antibiotic resistant bacteria express GBAP upon induction with isopropyl beta-D-thiogalactopyranoside (IPTG). Expression of GBAP in eukaryotic cells is achieved by infecting insect or mammalian cell lines with recombinant Autographica californica nuclear polyhedrosis virus (AcMNPV), commonly known as baculovirus. The nonessential polyhedrin gene of baculovirus is replaced with cDNA encoding GBAP by either homologous recombination or bacterial-mediated transposition involving transfer plasmid intermediates. Viral infectivity is maintained and the strong polyhedrin promoter drives high levels of cDNA transcription. Recombinant baculovirus is used to infect Spodoptera frugiperda (Sf9) insect cells in most cases, or human hepatocytes, in some cases. Infection of the latter requires additional genetic modifications to baculovirus. (See Engelhard, E.K. et al. (1994) Proc. Natl. Acad. Sci. USA 91:3224-3227; Sandig, V. et al. (1996) Hum. Gene Ther. 7:1937-1945.)

In most expression systems, GBAP is synthesized as a fusion protein with, e.g., glutathione S-transferase (GST) or a peptide epitope tag, such as FLAG or 6-His, permitting rapid, single-step, affinity-based purification of recombinant fusion protein from crude cell lysates. GST, a 26-kilodalton enzyme from Schistosoma japonicum, enables the purification of fusion proteins on immobilized glutathione under conditions that maintain protein activity and antigenicity (Amersham Pharmacia Biotech). Following purification, the GST moiety can be proteolytically cleaved from GBAP at specifically engineered sites. FLAG, an 8-amino acid peptide, enables immunoaffinity purification using commercially available monoclonal and polyclonal anti-FLAG antibodies (Eastman Kodak). 6-His, a stretch of six consecutive histidine residues, enables purification on metal-chelate resins (QIAGEN). Methods for protein expression and purification are discussed in Ausubel (1995, supra, ch. 10 and 16). Purified GBAP obtained by these methods can be used directly in the assays shown in Examples XI and XV.

XI. Demonstration of GBAP Activity

GTP-binding activity of GBAP is determined in an assay that measures the binding of GBAP

to α-P^{32}-labeled GTP. Purified GBAP is first blotted onto filters and rinsed in a suitable buffer. The filters are then incubated in buffer containing radiolabeled α-^{32}P-GTP. The filters are washed in buffer to remove unbound GTP and counted in a radioisotope counter. Non-specific binding is determined in an assay that contains a 100-fold excess of unlabeled GTP. The amount of specific binding is proportional to the activity of GBAP.

GTPase activity of GBAP is determined in an assay that measures the conversion of α-^{32}P-GTP to α-^{32}P-GDP. GBAP is incubated with α-^{32}P-GTP in buffer for an appropriate period of time, and the reaction is terminated by heating or acid precipitation followed by centrifugation. An aliquot of the supernatant is subjected to polyacrylamide gel electrophoresis (PAGE) to separate GDP and GTP together with unlabeled standards. The GDP spot is cut out and counted in a radioisotope counter. The amount of radioactivity recovered in GDP is proportional to GTPase activity of GBAP.

XII. Functional Assays

GBAP function is assessed by expressing the sequences encoding GBAP at physiologically elevated levels in mammalian cell culture systems. cDNA is subcloned into a mammalian expression vector containing a strong promoter that drives high levels of cDNA expression. Vectors of choice include pCMV SPORT plasmid (Life Technologies) and pCR3.1 plasmid (Invitrogen), both of which contain the cytomegalovirus promoter. 5-10 μg of recombinant vector are transiently transfected into a human cell line, for example, an endothelial or hematopoietic cell line, using either liposome formulations or electroporation. 1-2 μg of an additional plasmid containing sequences encoding a marker protein are co-transfected. Expression of a marker protein provides a means to distinguish transfected cells from nontransfected cells and is a reliable predictor of cDNA expression from the recombinant vector. Marker proteins of choice include, e.g., Green Fluorescent Protein (GFP; Clontech), CD64, or a CD64-GFP fusion protein. Flow cytometry (FCM), an automated, laser optics-based technique, is used to identify transfected cells expressing GFP or CD64-GFP and to evaluate the apoptotic state of the cells and other cellular properties. FCM detects and quantifies the uptake of fluorescent molecules that diagnose events preceding or coincident with cell death. These events include changes in nuclear DNA content as measured by staining of DNA with propidium iodide; changes in cell size and granularity as measured by forward light scatter and 90 degree side light scatter; down-regulation of DNA synthesis as measured by decrease in bromodeoxyuridine uptake; alterations in expression of cell surface and intracellular proteins as measured by reactivity with specific antibodies; and alterations in plasma membrane composition as measured by the binding of fluorescein-conjugated Annexin V protein to the cell surface. Methods in flow cytometry are discussed in Ormerod, M.G. (1994) Flow Cytometry, Oxford, New York NY.

The influence of GBAP on gene expression can be assessed using highly purified populations of cells transfected with sequences encoding GBAP and either CD64 or CD64-GFP. CD64 and CD64-

GFP are expressed on the surface of transfected cells and bind to conserved regions of human immunoglobulin G (IgG). Transfected cells are efficiently separated from nontransfected cells using magnetic beads coated with either human IgG or antibody against CD64 (DYNAL, Lake Success NY). mRNA can be purified from the cells using methods well known by those of skill in the art. Expression of mRNA encoding GBAP and other genes of interest can be analyzed by northern analysis or microarray techniques.

XIII. Production of GBAP Specific Antibodies

GBAP substantially purified using polyacrylamide gel electrophoresis (PAGE; see, e.g., Harrington, M.G. (1990) Methods Enzymol. 182:488-495), or other purification techniques, is used to immunize rabbits and to produce antibodies using standard protocols.

Alternatively, the GBAP amino acid sequence is analyzed using LASERGENE software (DNASTAR) to determine regions of high immunogenicity, and a corresponding oligopeptide is synthesized and used to raise antibodies by means known to those of skill in the art. Methods for selection of appropriate epitopes, such as those near the C-terminus or in hydrophilic regions are well described in the art. (See, e.g., Ausubel, 1995, supra, ch. 11.)

Typically, oligopeptides of about 15 residues in length are synthesized using an ABI 431A peptide synthesizer (PE Biosystems) using FMOC chemistry and coupled to KLH (Sigma-Aldrich, St. Louis MO) by reaction with N-maleimidobenzoyl-N-hydroxysuccinimide ester (MBS) to increase immunogenicity. (See, e.g., Ausubel, 1995, supra.) Rabbits are immunized with the oligopeptide-KLH complex in complete Freund's adjuvant. Resulting antisera are tested for antipeptide and anti-GBAP activity by, for example, binding the peptide or GBAP to a substrate, blocking with 1% BSA, reacting with rabbit antisera, washing, and reacting with radio-iodinated goat anti-rabbit IgG.

XIV. Purification of Naturally Occurring GBAP Using Specific Antibodies

Naturally occurring or recombinant GBAP is substantially purified by immunoaffinity chromatography using antibodies specific for GBAP. An immunoaffinity column is constructed by covalently coupling anti-GBAP antibody to an activated chromatographic resin, such as CNBr-activated SEPHAROSE (Amersham Pharmacia Biotech). After the coupling, the resin is blocked and washed according to the manufacturer's instructions.

Media containing GBAP are passed over the immunoaffinity column, and the column is washed under conditions that allow the preferential absorbance of GBAP (e.g., high ionic strength buffers in the presence of detergent). The column is eluted under conditions that disrupt antibody/GBAP binding (e.g., a buffer of pH 2 to pH 3, or a high concentration of a chaotrope, such as urea or thiocyanate ion), and GBAP is collected.

XV. Identification of Molecules Which Interact with GBAP

GBAP, or biologically active fragments thereof, are labeled with ^{125}I Bolton-Hunter reagent.

(See, e.g., Bolton A.E. and W.M. Hunter (1973) Biochem. J. 133:529-539.) Candidate molecules previously arrayed in the wells of a multi-well plate are incubated with the labeled GBAP, washed, and any wells with labeled GBAP complex are assayed. Data obtained using different concentrations of GBAP are used to calculate values for the number, affinity, and association of GBAP with the candidate molecules.

Alternatively, molecules interacting with GBAP are analyzed using the yeast two-hybrid system as described in Fields, S. and O. Song (1989, Nature 340:245-246), or using commercially available kits based on the two-hybrid system, such as the MATCHMAKER system (Clontech).

GBAP may also be used in the PATHCALLING process (CuraGen Corp., New Haven CT) which employs the yeast two-hybrid system in a high-throughput manner to determine all interactions between the proteins encoded by two large libraries of genes (Nandabalan, K. et al. (2000) U.S. Patent No. 6,057,101).

Various modifications and variations of the described methods and systems of the invention will be apparent to those skilled in the art without departing from the scope and spirit of the invention. Although the invention has been described in connection with certain embodiments, it should be understood that the invention as claimed should not be unduly limited to such specific embodiments. Indeed, various modifications of the described modes for carrying out the invention which are obvious to those skilled in molecular biology or related fields are intended to be within the scope of the following claims.

Table 1

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
1	67	1405545	LATRTUT02	1405545F6 (LATRTUT02), 1405545H1 (LATRTUT02), 2926327F7 (TLYMNOT04), 2926327T6 (TLYMNOT04)
2	68	1451265	PENITUT01	700515X14 (SYNORAT03), 758541H1 (BRAITUT02), 1348685F6 (PROSNOT11), 1451265H1 (PENITUT01), 1872777F6 (LEUKNOT02)
3	69	1556311	BLADTUT04	1556311H1 (BLADTUT04), 3221281T6 (COLNNON03), 3350311F6 (BRAITUT24), SBFA02256F1, SBFA01440F1, SBFA01098F1, SBFA04741F1
4	70	1901373	BLADTUT06	758057H1 (BRAITUT02), 1255886H1 (MENITUT03), 1887731X12C1 (BLADTUT07), 1901373H1 (BLADTUT06), 2866863H1 (KIDNNOT20), 3090943H1 (BRSTNOT19), 3215237H1 (TESTNOT07), 3719233H1 (PENCNOT10), 4319601H1 (BRADDIT02)
5	71	2367767	ADRENOT07	1331124F1 (PANCNOT07), 2367767H1 (ADRENOT07), 2367779F6 (ADRENOT07) 2782232F6 (BRSTNOT13), 3079286H2 (BRAIUNT01), 3584043T6 (293TF4T01), 4994696H1 (LIVRTUT11)
6	72	3090433	BRSTNOT19	312565H1 (LUNGNOT02), 841829R6 (PROSTUT05), 1340809H1 (COLNTUT03), 1842057H1 (COLNNOT07), 2693513F6 (LUNGNOT23), 3090433H1 (BRSTNOT19), 4895874H1 (LIVRTUT12)
7	73	3800591	SPLNNOT12	554715F1 (SCORNOT01), 882035X23 (THYRNOT02), 3042234F7 (BRSTNOT16), 3630695H1 (COLNNOT38), 3800591H1 (SPLNNOT12), 4975447H1 (HELATXT03)
8	74	5308471	MONOTXT02	790680R1 (PROSTUT03), 870507R1 (LUNGAST01), 948177R1 (PANCNOT05), 1682469T7 (PROSNOT15), 2897215H1 (KIDNTUT14), 5308471H1 (MONOTXT02)
9	75	5324322	FIBPFEN06	1001977R1 (BRSTNOT03), 1312045F1 (COLNFET02), 1334040F2 (COLNNOT13), 1488082F6 (UCMCL5T01), 1570077F1 (UTRSNOT05), 1929845H1 (COLNTUT03), 2306061H1 (NGANNOT01), 3127730F7 (LUNGTUT12), 3494367H1 (ADRETUT07), 3578924H1 (293TF3T01), 4619513H1 (ENDVNOT01), 4932823H1 (BRSTTUT20), 5324322H1 (FIBPFEN06)
10	76	067184	HUVESTB01	067184H1 (HUVESTB01), 067184R1 (HUVESTB01), 067184X12 (HUVESTB01), 067184X23C1 (HUVESTB01), 067184X29C1 (HUVESTB01), 968551H1 (BRSTNOT05), 2611874T6 (LUNGTUT10)
11	77	722896	SYNOOAT01	722896H1 (SYNOOAT01), 722896X19C1 (SYNOOAT01), 1433775T1 (BEPINON01), 1477633T6 (CORPNOT02), 2676923F6 (KIDNNOT19), 3230945H1 (COTRNOT01), 3389989H1 (LUNGTUT17)
12	78	1571739	UTRSNOT05	1571739H1 (UTRSNOT05), 1571739X12R1 (UTRSNOT05), 2799982H1 (PENCNOT01), 4059114F6 (BRAINOT21)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
13	79	1739479	HIPONON01	511157H1 (MPHGNOT03), 511157T6 (MPHGNOT03), 1739479H1 (HIPONON01), 2092446T6 (PANCNOT04), 3880948F6 (SPLNNOT11)
14	80	1999147	BRSTTUT03	1339243T6 (COLNTUT03), 1999147H1 (BRSTTUT03), 2094940X11F1 (BRAITUT02), 2670959T6 (ESOGTUT02), 3297709H1 (TLYJINT01), 3396927H1 (UTRSNOT16), SCBA00828V1, SCBA00615V1, SCBA04422V1, SCBA04646V1, SCBA01715V1, 5544151H1 (TESTNOC01)
15	81	2182085	SININOT01	767764R6 (LUNGNOT04), 1655010F6 (PROSTUT08), 1701703T6 (BLADTUT05), 1871360F6 (SKINBIT01), 2081835F6 (UTRSNOT08), 2411644H1 (BSTMNON02)
16	82	2216640	SINTFET03	489759H1 (HNT2AGT01), 2057454T6 (BEPINOT01), 2097739H1 (BRAITUT02), 2216640H1 (SINTFET03), 2325135H1 (OVARNOT02), 2361273R6 (LUNGFET05), 2667958H1 (ESOGTUT02), 3462348H1 (293TF2T01), 3478754H1 (OVARNOT11), 4163069F6 (BRSTNOT32)
17	83	2417361	HNT3AZT01	1394742F1 (THYRNOT03), 2417361F6 (HNT3AZT01), 2417361H1 (HNT3AZT01)
18	84	2454384	ENDANOT01	2454384H1 (ENDANOT01), 2454384T6 (ENDANOT01), 2589653T6 (LUNGNOT22), 2643485F6 (LUNGTUT08), 2723048H1 (LUNGTUT10), 3130367H1 (LUNGTUT12)
19	85	2610262	LUNGTUT08	1226946R6 (COLNNOT01), 1226946T6 (COLNNOT01), 2610262F6 (LUNGTUT08), 2610262H1 (LUNGTUT08)
20	86	2700075	OVARTUT10	604199R1 (BRSTTUT01), 1225126R1 (COLNTUT02), 1923323R6 (BRSTTUT01), 2301778R6 (BRSTNOT05), 2506882F6 (CONUTUT01), 2700075F6 (OVARTUT10), 2700075H1 (OVARTUT10), 2744960F6 (LUNGTUT11), 2833994F6 (TLYMNOT03), 2915413H1 (THYMFET03), 3647274H1 (ENDINOT01)
21	87	2786701	BRSTNOT13	754370R1 (BRAITUT02), 1426163R6 (BEPINON01), 1850667F6 (LUNGFET03), 1923562R6 (BRSTTUT01), 2215161F6 (SINTFET03), 2215161T6 (SINTFET03), 2498589H1 (ADRETUT05), 2991672F6 (KIDNFET02), 3028991H1 (HEARFET02), 3729514H1 (SMCCNON03), 5065467H1 (ARTFTDT01)
22	88	3068538	UTRSNOR01	908465R2 (COLNNOT09), 957130R6 (KIDNNOT05), 1301520F6 (BRSTNOT07), 1580628H1 (DUODNOT01), 2631247F6 (COLNTUT15), 3068538H1 (UTRSNOR01), 3532286T6 (KIDNNOT25)
23	89	5159072	BRSTTMT02	412241R1 (BRSTNOT01), 660435H1 (BRAINOT03), 881160H1 (THYRNOT02), 1304119F6 (PLACNOT02), 1324073F1 (LPARNOT02), 2520427H1 (BRAITUT21), 5159072H1 (BRSTTMT02)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
24	90	5519057	LIVRDIR01	066809H1 (HUVESTB01), 3279230H1 (STOMFET02), 5370305F6 (BRAINOT22), 5508943F6 (BRADDIR01), 5508943R6 (BRADDIR01), 5519057H1 (LIVRDIR01)
25	91	035379	HUVENOB01	035379H1 (HUVENOB01), 035379X11 (HUVENOB01), 035379X12 (HUVENOB01), 035379X13 (HUVENOB01), 035379X311D1 (HUVENOB01), 112161R1 (PITUNOT01), 1922877R6 (BRSTTUT01), 2133108F6 (ENDCNOT01), 3107232H1 (BRSTTUT15), 4798135H1 (LIVRTUT09), SCHA01519V1, g1802757
26	92	275354	TESTNOT03	275354H1 (TESTNOT03), 275354X1 (TESTNOT03), 1663122T6 (BRSTNOT09), 2104284R6 (BRAITUT02), 2738788T6 (OVARNOT09), 3584082T6 (293TF4T01), SCGA07807V1
27	93	311658	LUNGNOT02	207452X12 (SPLNNOT02), 238306X85F1 (SINTNOT02), 264489H1 (HNT2AGT01), 311658H1 (LUNGNOT02), 1292829F6 (PGANNOT03), 1298271F1 (BRSTNOT07), 1488285H1 (UCMCL5T01), 2555757H1 (THYMNOT03), 2665984F6 (ADRENOT08), 2665984T6 (ADRENOT08), 3079209H1 (BRAIUNT01)
28	94	1251632	LUNGFET03	1251632H1 (LUNGFET03), 1251632X11 (LUNGFET03), 1251632X13 (LUNGFET03), 1316814T1 (BLADTUT02), 1384212F1 (BRAITUT08), 1711274F6 (PROSNOT16), 3128230H1 (LUNGTUT12), 4819602H1 (PROSTUT17), SZZZ00620R1
29	95	1331955	PANCNOT07	1363667X12 (LUNGNOT12), 1363667X13 (LUNGNOT12), SBBA01489F1, SBBA01528F1
30	96	1412614	BRAINOT12	1412614F6 (BRAINOT12), 1412614H1 (BRAINOT12), 2278130H1 (PROSNON01), 2278130T6 (PROSNON01), 5105388T6 (PROSTUS19)
31	97	1750781	LIVRTUT01	452712T6 (TLYMNOT02), 483862R6 (HNT2RAT01), 777729R6 (COLNNOT05), 1394724F1 (THYRNOT03), 1652134F6 (PROSTUT08), 1750781F6 (LIVRTUT01), 1750781H1 (LIVRTUT01), 1750781X305F1 (LIVRTUT01), 1750781X307D2 (LIVRTUT01), 3221477H1 (COLNNON03), SCHA02984V1, SXAA02156D1, SXAA00802D1
32	98	1821658	GBLATUT01	909674H1 (STOMNOT02), 1579095F1 (DUODNOT01), 1821658H1 (GBLATUT01), 1821658T6 (GBLATUT01), 2508922F6 (CONUTUT01), 2584263H1 (BRAITUT22), 5571821H1 (TLYMNOT08)
33	99	1872574	LEUKNOT02	305990F1 (HEARNOT01), 908252R2 (COLNNOT09), 1872574H1 (LEUKNOT02), 2051868F6 (LIVRFET02), 2285632R6 (BRAINON01), 3181732F6 (TLYJNOT01), 3285854F6 (HEAONOT05), 3332012H1 (BRAIFET01), SBWA02751V1, SBWA02849V1, SBWA04744V1, SBWA00180V1

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
34	100	2590967	LUNGNOT22	1340471F6 (COLNTUT03), 2590967F6 (LUNGNOT22), 2590967H1 (LUNGNOT22), 2771160F6 (COLANOT02), 3150287R6 (ADRENON04)
35	101	2824491	ADRETUT06	1381834X14 (BRAITUT08), 1381834X16 (BRAITUT08), 1381834X17 (BRAITUT08), 1381834X31 (BRAITUT08), 1972345F6 (UCMCL5T01), 2824491H1 (ADRETUT06), 3413970H1 (PTHYNOT04)
36	102	2825460	ADRETUT06	870873R6 (LUNGAST01), 1440326F1 (THYRNOT03), 2825460H1 (ADRETUT06), 2825460T6 (ADRETUT06), 4154518H1 (MUSLTMT01), 5068209H1 (PANCNOT23), SBLA03097F1
37	103	2871116	THYRNOT10	357664R6 (PROSNOT01), 1419595F1 (KIDNNOT09), 1419595T1 (KIDNNOT09), 1577877F6 (LNODNOT03), 1577877T1 (LNODNOT03), 2767635H1 (COLANOT02), 2871116F6 (THYRNOT10), 2871116H1 (THYRNOT10), 4650546H1 (PROSTUT20), SBHA03160F1, SBHA02613F1, SBHA02703F1
38	104	2942212	CONNTUT05	1270807H1 (TESTTUT02), 1270807X301D1 (TESTTUT02), 1270807X309D2 (TESTTUT02), 2942212H2 (CONNTUT05), g1924758
39	105	3685151	HEAANOT01	860843R1 (BRAITUT03), 1932207F6 (COLNNOT16), 1932207T6 (COLNNOT16), 2210580F6 (SINTFET03), 3043060H1 (HEAANOT01), 3685151H1 (HEAANOT01), 4960825H1 (TLYMNOT05)
40	106	4881515	UTRMTMT01	925415R1 (BRAINOT04), 1337450F6 (COLNNOT13), 1961288R6 (BRSTNOT04), 3581069H1 (293TF3T01), 3583842T6 (293TF4T01), 4881515H1 (UTRMTMT01), 5488514H1 (DRGTNON04), g1156606
41	107	5324681	FIBPFEN06	2455960T6 (ENDANOT01), 2458281F6 (ENDANOT01), 3834084F6 (PANCNOT17), 4046332H1 (LUNGNOT35), 5324681H1 (FIBPFEN06), g1733388, g1522074
42	108	5387651	BRAINOT19	810934T1 (LUNGNOT04), 822997R1 (KERANOT02), 1282647F1 (COLNNOT16), 1282647T1 (COLNNOT16), 1571430T6 (UTRSNOT05), 2208839F6 (SINTFET03), 2844787H1 (DRGLNOT01), 2908748H1 (THYMNOT05), 5387651H1 (BRAINOT19)
43	109	5595679	COLCDIT03	044292R6 (TBLYNOT01), 826501R1 (PROSNOT06), 1251632X12 (LUNGFET03), 1303934F1 (PLACNOT02), 1316814F1 (BLADTUT02), 1339567T1 (COLNTUT03), 2806159H1 (BLADTUT08), 2837021H1 (TLYMNOT03), 3037493H1 (BRSTNOT16), 3119883H1 (LUNGTUT13), 3395946H1 (LUNGNOT28), 3748742H1 (UTRSNOT18)
44	110	5782457	BRAXNOT03	532593R6 (BRAINOT03), 532593T6 (BRAINOT03), 5782457H1 (BRAXNOT03)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
45	111	760677	BRAITUT02	745006X13 (BRAITUT01), 760677H1 (BRAITUT02), 760677X19 (BRAITUT02), 763135X12 (BRAITUT02), 946075H1 (RATRNOT02), 953938H1 (SCORNON01)
46	112	1348567	PROSNOT11	1348567H1 (PROSNOT11), 1505075F6 (BRAITUT07), 1620627F6 (BRAITUT13), 2069105F6 (ISLTNOT01), 2417901F6 (HNT3AZT01), 2494683H1 (ADRETUT05), 3320166H1 (PROSBPT03)
47	113	1751354	LIVRTUT01	029909F1 (SPLNFET01), 029909R1 (SPLNFET01), 512371H1 (MPHGNOT03), 1439362F6 (PANCNOT08), 1751354F6 (LIVRTUT01), 1751354H1 (LIVRTUT01), 1900168F6 (BLADTUT06)
48	114	1976780	PANCTUT02	001347H1 (U937NOT01), 1755035X307D2 (LIVRTUT01), 1976780H1 (PANCTUT02), 2798389H1 (NPOLNOT01), 4050076H1 (SINTNOT18), 4228943H1 (BRAMDIT01), 4291877H1 (BRABDIR01), 5514957H1 (BRADDIR01), SCHA04173V1, SCHA02986V1, SCHA01162V1, SCIA02096V1
49	115	2048234	LIVRFET02	1553355F6 (BLADTUT04), 1929455F6 (COLNTUT03), 2048234H1 (LIVRFET02), 2699864T6 (OVARTUT10)
50	116	2111754	BRAITUT03	1335055F6 (COLNNOT13), 2105233R6 (BRAITUT03), 2111754H1 (BRAITUT03), 2111754R6 (BRAITUT03), 3706377H1 (PENCNOT07)
51	117	2123286	BRSTNOT07	411359F1 (BRSTNOT01), 411359R1 (BRSTNOT01), 708105R6 (SYNORAT04), 1322780F6 (BLADNOT04), 2123286H1 (BRSTNOT07), 2719651F6 (LUNGTUT10), 2880143F6 (UTRSTUT05), 3206153F6 (PENCNOT03), 3210501F6 (BLADNOT08), 3346625F6 (BRAITUT24), 3489118H1 (EPIGNOT01), 3605764H1 (LUNGNOT30), 4242993H1 (SYNWDIT01), 5089472H1 (UTRSTMR01)
52	118	2477507	SMCANOT01	488096H1 (HNT2AGT01), 1672690F6 (BLADNOT05), 1802830F6 (COLNNOT27), 1818538H1 (PROSNOT20), 2171841H1 (ENDCNOT03), 2477507H1 (SMCANOT01), 3434030F6 (PENCNOT05)
53	119	2759119	THP1AZS08	496782H1 (HNT2NOT01), 1251166H1 (LUNGFET03), 1289067F1 (BRAINOT11), 1295658T6 (PGANNOT03), 1510901F1 (LUNGNOT14), 1531583F1 (SPLNNOT04), 1533488F1 (SPLNNOT04), 1817447H1 (PROSNOT20), 2154846F6 (BRAINOT09), 2468875H1 (THYRNOT08), 2498852F6 (ADRETUT05), 2506652F6 (CONUTUT01), 2630812F6 (COLNTUT15), 2759119H1 (THP1AZS08), 2991227H1 (KIDNFET02), 3036646F6 (PENCNOT02), 3213032H1 (BLADNOT08)
54	120	2823818	ADRETUT06	618671R6 (PGANNOT01), 2823818H1 (ADRETUT06), 2950988F6 (KIDNFET01), g1679455

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
55	121	2859730	SININOT03	103901X6 (BMARNOT02), 510695H1 (MPHGNOT03), 1452088H1 (PENITUT01), 1527095F6 (UCMCL5T01), 2285371H1 (BRAINON01), 2843029H1 (DRGLNOT01), 2859730H1 (SININOT03)
56	122	2861155	SININOT03	875215T1 (LUNGAST01), 999673H1 (KIDNTUT01), 1425091R6 (BEPINON01), 2861155F6 (SININOT03), 2861155H1 (SININOT03), 2901915F6 (DRGCNOT01), 3621947H2 (ENDANOT03)
57	123	3002667	TLYMNOT06	227882F1 (PANCNOT01), 227882R1 (PANCNOT01), 260725H1 (HNT2RAT01), 1432542R1 (BEPINON01), 2474761F6 (SMCANOT01), 3002667H1 (TLYMNOT06), 3188977H1 (THYMNON04), 3461163H1 (293TF1T01), 4860339F6 (PROSTUT09)
58	124	3043734	HEAANOT01	3043734H1 (HEAANOT01), 3043734T6 (HEAANOT01), 3209823H1 (BLADNOT08), 5277071H1 (MUSLNOT01)
59	125	3294893	TLYJINT01	389234H1 (THYMNOT02), 1242886H1 (LUNGNOT03), 1539958T1 (SINTTUT01), 1870567H1 (SKINBIT01), 2069284F6 (ISLTNOT01), 2280217R6 (PROSNON01), 2353465T6 (LUNGNOT20), 2798990F6 (NPOLNOT01), 3180440H1 (TLYJNOT01), 3294893H1 (TLYJINT01), 3816962H1 (TONSNOT03), 5039889H2 (COLHTUT01), 5118831H1 (SMCBUNT01)
60	126	3349052	BRAITUT24	731775H1 (LUNGNOT03), 1449575H1 (PLACNOT02), 1899442F6 (BLADTUT06), 1967162T6 (BRSTNOT04), 2630025F6 (COLNTUT15), 2717821H1 (THYRNOT09), 3180478T6 (TLYJNOT01), 3349052H1 (BRAITUT24), 4523961F6 (HNT2TXT01), 5565623H1 (TLYMNOT08), 6141909H1 (BMARTXT03)
61	127	3357264	PROSTUT16	2378150F6 (ISLTNOT01), 2378150X304B1 (ISLTNOT01), 2378150X304D1 (ISLTNOT01), 2807493F6 (BLADTUT08), 2881251F6 (UTRSTUT05), 3357264F6 (PROSTUT16), 3357264H1 (PROSTUT16), 3593272H1 (293TF5T01), 4163652T6 (BRSTNOT32), 4821588F6 (PROSTUT17), 4872125H1 (COLDNOT01)
62	128	3576329	BRONNOT01	1444072F6 (THYRNOT03), 1649584T6 (PROSTUT09), 1720770X15C1 (BLADNOT06), 1720770X16C1 (BLADNOT06), 2204612F6 (SPLNFET02), 3576329H1 (BRONNOT01), SAFC01083F1
63	129	3805550	BLADTUT03	1416364F6 (BRAINOT12), 1553473H1 (BLADTUT04), 3232384H1 (COLNUCT03), 3287257H1 (HEAONOT05), 3539473H1 (SEMVNOT04), 3805550H1 (BLADTUT03)

Table 1 (cont.)

Protein SEQ ID NO:	Nucleotide SEQ ID NO:	Clone ID	Library	Fragments
64	130	4546403	COLXTDT01	1687704F6 (PROSTUT10), 1962744R6 (BRSTNOT04), 2674742F6 (KIDNNOT19), 4546403H1 (COLXTDT01), 4632828T6 (GBLADIT02)
65	131	4767318	BRATNOT02	134566R1 (BMARNOT02), 549352R1 (BEPINOT01), 1819757T6 (GBLATUT01), 2863295H1 (KIDNNOT20), 4767318H1 (BRATNOT02), SBLA03778F1, g3737930
66	132	4834527	BRAWNOT01	859906X38C1 (BRAITUT03), 1231225H1 (BRAITUT01), 1393681T6 (THYRNOT03), 1416996F6 (BRAINOT12), 2422475H1 (SCORNON02), 3999137R6 (HNT2AZS07), 4834527F6 (BRAWNOT01), 4834527H1 (BRAWNOT01), 5691642H1 (BRAUNOT02)

Table 2

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
1	269	S59 T71 T146 T211 T73 S127 T133 S216	N12	GTP-binding protein: D79-M234, Y80-C239 ATP/GTP binding site (P-loop): G102-S109	GTP-binding protein; CgpA [Caulobacter crescentus] g3820578	BLAST-Genbank BLAST-DOMO MOTIFS
2	428	S59 S188 S200 S284 S367 S381 T399 T29 T193 T288 T354 S419		Beta transducin family, G-beta repeats: T269-L315 , F261-D293 L280-V294 , V185-V199 Signal peptide: M1-A35		ProfileScan MOTIFS BLIMPS-PRINTS HMMER-PFAM SPScan
3	562	S151 S152 T443 T444 S33 S104 S126 S127 S135 S216 S239 T350 T383 S450 T481 S146 T223 S287 S356 T434 T470 Y501	N125 N354 N445		Ras inhibitor [Homo sapiens] g190895	BLAST-Genbank
4	229	T108 S153 S9 S160 S215 T219 T142 S180	N111 N140 N198	ATP/GTP-binding site: G28-S35 Ras family: K23-T219 Ras transforming protein: V22-M43, A63-S85, P124-A137, L156-A178, D102-S145, K150-S180	Small GTP binding protein [Saccharomyces cerevisiae] g1171484	BLAST-Genbank MOTIFS HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-DOMO
5	360	T108 S360 S115 T217 T264 S295 S296 S35 S52 S160 S174 T206 T249	N149 N287 N327 N351	WD domain, G-beta repeats: M1-T64, M27-K41, F274-K306	Similar to WD domain, G-beta repeat protein [C. elegans] g3880929	BLAST-Genbank HMMER-PFAM ProfileScan BLIMPS-PRINTS
6	460	T18 T107 T123 S149 S199 S280 S336 S369 S71 T106 S387 Y302 Y400	N270 N350	Signal peptide: M1-A57	Rabin3 [Rattus norvegicus] g624225	BLAST-Genbank SPScan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
7	239	S234 S25 T47 T52 S98 T190 T206 S236 S223	N188	Phosducin: L20-I179, S25-I179, E30-D239	Phosducin-like protein [Homo sapiens] g4104075	BLAST-Genbank BLAST-PRODOM BLAST-DOMO
8	334	T225 T235 S260 T4 S45 S63 S133 S162 S193 T279 T308		ATP/GTP-binding site (P-loop): G150-S157 GTP1/OBG family: L75-D89, I146-Q166 G-protein, alpha subunit: I79-L87	GTP-binding protein homolog [L. braziliensis] g2570231	BLAST-Genbank MOTIFS BLIMPS-BLOCKS BLIMPS-PRINTS
9	341	S91 T122 S185 T199 T228 S65 T85 S323		Signal peptide: M1-A61 WD domain,G-beta repeats: L164-D196, C173-P217, V183-L197, S185-W195	Putative WD-40 repeat protein [Arabidopsis thaliana] g4191773	SPScan BLAST-Genbank MOTIFS ProfileScan HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS
10	513	T29 T72 T109 S124 S136 S215 T341 T481 T501 S65 T245 T330 S338 T372 T386 S437 S451 T473 Y228 Y254	N242 N417	Beta-transducin family, G-beta repeats: F345-N377, K210-N242, E303-G335, S366-W376, N353-V400, L229-F243, I364-M378	Similar to WD domain G-beta repeats protein [C. elegans] g3875246	BLAST-Genbank MOTIFS HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS ProfileScan
11	186	T61 S80 S107 S163 S31 T66 S183	N64 N148	ARF-family: N6-S186, P51-S90, M95-L149 GTP-binding, SAR1 protein: F78-K103, I123-I144 ATP/GTP binding site (P-loop): G27-T34	Similar to ADP-ribosylation factor [C. elegans] g3881189	BLAST-Genbank HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS MOTIFS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
12	204	S184 S203 S34 S152 T14 T20 T25 T62 S86		Ras family: K5-M189 Ras transforming protein: M1-E150, V4-T25, V113-L126 ATP/GTP binding site (P-loop): G10-S17	Ras-like protein, rit [Mus musculus] g1656005	BLAST-Genbank HMMER-PFAM BLIMPS-PRINTS BLAST-DOMO MOTIFS
13	100	S31 S46 T52 T61 S84 S4 S26 S27 T86		Beta-transducin, WD repeats: L81-M95, V70-S100, M1-S100	Similar to beta-transducin [C. elegans] g3875373; Alzheimer's disease protein [Homo sapiens] GeneSeq W21578	BLAST-Genbank MOTIFS BLIMPS-BLOCKS ProfileScan BLIMPS-PRINTS BLAST-PRODOM
14	795	T569 S776 S54 S188 S201 T248 T249 T298 S306 S368 T422 S466 T561 S586 S625 S678 T731 S777 S13 T42 S120 T134 T174 S213 S254 T266 S391 S415 S588 S620 S694 T742	N52 N421 N559 N585 N708	WD domain, G-beta repeats: L108-L139, L147-K179, T168-W178, Y227-K259, L126-N140, M166-A180	Phospholipase A2-activating protein [Rattus Norvegicus] g1017706	BLAST-Genbank BLAST-PRODOM BLAST-DOMO HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS
15	393	S48 S61 T143 T334 T148 T200 S208 T212 T245 S266 S325	N182 N197	WD domain, G-beta repeats: L121-A153, L357-R389, P322-F369, L140-S154	Putative WD-repeat protein [Arabidopsis thaliana] g4263521	BLAST-Genbank HMMER-PFAM ProfileScan BLIMPS-PRINTS
16	485	S31 S108 S222 S321 S346 S357 T84 T125 T137 T151 T187 S227 T268 S395 T403 S409 T437 Y92 Y261		Beta-transducin, WD repeats: L129-L143, V219-T233, S262-W272, V387-G401, L429-V443, L452-G468	Notchless protein [Xenopus laevis] g3687833	BLAST-Genbank MOTIFS HMMER-PFAM ProfileScan BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-DOMO BLAST-PRODOM

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
17	199	T32 T91 S177 T56 S153 S186 Y149		ATP/GTP-binding site (P-loop): G15-T22 Transforming protein, p21: L9-H30, T32-K48, I50-S72, Q115-L128, Y149-A171 Ras protein: K5-E151	Rab7 [Mus musculus] g1050551	BLAST-Genbank MOTIFS BLIMPS-PRINTS BLAST-PRODOM BLAST-DOMO
18	163	T18 T46 S120 S5 T151 T83 S125	N81 N159		Rhotekin [Mus musculus] g1293145	BLAST-Genbank
19	290	S56 S84 T234 S41 T91 T132 T234 T11 T47 T80 T194	N89 N188	Beta-transducin, WD-repeats: S41-W51, F195-D227, L238-N270, L214-I228, L257-M271, T203-S249	Similar to beta-transducin; [C. elegans] g3875373; Alzheimer's disease protein [Homo sapiens] GeneSeq W21578	BLAST-Genbank MOTIFS HMMER-PFAM BLIMPS-BLOCKS ProfileScan BLIMPS-PRINTS BLAST-PRODOM
20	705	T277 T364 S393 S448 S479 S483 T554 T568 S586 S239 S250 T374 S379 T398 S485 T528	N274	Beta-transducin, WD-repeats: L390-L404, L370-D403, L413-R445	Similar to WD domain G-beta repeat prot. [C. elegans] g3880340; 70kD tumor-specific antigen [R. norvegicus] g2505957	BLAST-Genbank HMMER-PFAM BLAST-PRODOM BLAST-DOMO BLIMPS-BLOCKS BLIMPS-PRINTS MOTIFS
21	454	T426 S451 S28 S51 T81 T89 T166 S214 T241 S264 T305 S343 S185 T193 S421	N58	ATP/GTP-binding site (P-loop): G73-S80 Cell division control protein: V47-P240	Similar to Drosophila melanogaster septin (sep2) [Homo sapiens] g1503988	BLAST-Genbank BLAST-PRODOM BLAST-DOMO MOTIFS
22	433	S169 T239 T292 S309 S382 S129 S297 Y60 Y101 Y315	N338	Protein GTPase activating protein: L8-S169 PH domain: Y138-Q355, Q191-I351, P210-E375	RhoGAP protein [Homo sapiens] g312212	BLAST-Genbank BLAST-PRODOM BLAST-DOMO

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
23	406	T83 S143 S303 T75 T115 T126 T211 S216 T289 T315 Y247	N184 N401 N402		Rab 9 effector, P40 [Homo sapiens] g2217970	BLAST-Genbank
24	229	S7 S127 T50 S178		ATP/GTP-binding site (P-loop): G40-T47 Ras family: K35-L217 Transforming protein, p21: F34-A55, R57-R73, V75-K97, N139-L152	Rab GTPase, Rab33B [Mus musculus] g2516239	BLAST-Genbank MOTIFS HMMER-PFAM BLIMPS-PRINTS BLAST-DOMO
25	670	T28 T45 S69 S3 S108 T277 S406 S6 T52 T82 S91 S102 S126 S609 S158 S197 T213 S217 T281 S323 S416 T419 T428 T474 S496 T540 S624 T664	N343	G-beta WD repeat domain: F386-D424, L411-T425, Y429-D465, L469-D504, L510-D545, L549-D585, K589-S629, M633-T669 Beta-transducin Trp-Asp repeats signature: C401-I447	Beta transducin-like protein [Podospora anserina] g607003	BLAST-GenBank BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS ProfileScan
26	445	T17 T48 T126 T160 T293 T364 T97 T132 S201 S217 S305 T322 S357 S434 Y339	N46 N95 N355	G-beta WD repeat domain: L62-N95, V82-L96, F124-M138, F297-V311 Beta-transducin Trp-Asp repeats signature: S316-A356 SOF1 protein, WD repeat: D129-V277, F309-V444	Beta-transducin [Schizosaccharomyces pombe] g3393019	BLAST-GenBank BLAST-DOMO BLAST-PRODOM BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS
27	236	S24 S60 S86 T181 S117 S140		GYP7, GTPase activating protein: M1-I155	GTPase activating protein [Yarrowia lipolytica] g2370595	BLAST-GenBank BLAST-PRODOM MOTIFS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
28	498	S97 T158 S247 S281 S425 S468 S494 T84 S176 T355 T474 Y239		G-beta WD repeat domain: L188-Q220, L446-G479, M466-P480 Beta-transducin Trp-Asp repeats signature: F200-A245	Similarity to guanine nucleotide binding protein [Caenorhabditis elegans] g3878300	BLAST-GenBank BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS ProfileScan
29	334	S63 S104 S148 S189 T208 S276 S50 T110 S118 T124 S152 T160 T237 T326	N265	G-beta WD repeat domain: L41-G73, I83-D115, L102-V116, L125-D157, L167-D199, I210-D242 Beta-transducin Trp-Asp repeats signature: S49-A308 Signal peptide: M1-A47	Similar to guanine nucleotide binding protein [Caenorhabditis elegans] g3874290	BLAST-GenBank BLAST-PRODOM BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS ProfileScan SPScan
30	292	S102 T145 S188 S52 T89 S204 S222 S283	N209	Protein with WD repeat: P7-W129 Signal peptide: M1-S68	F-box protein FBX16 [Mus musculus] g6456114	BLAST-PRODOM BLAST-GenBank MOTIFS SPScan
31	588	T184 T76 T137 S139 T161 T174 T183 S285 T351 T375 S432 T473 S488 S213 T265 S389 S394 T412 T546	N159	G-beta WD repeat domain: A293-E331, C337-T375, Y379-D417, I404-L418, E460-D497, T506-S543, G547-A586 Beta-transducin Trp-Asp repeats signature: A308-E354, L393-Q441	TipD (sequence similarity to Beta-transducin family) [Dictyostelium discoideum] g2407788	BLAST-GenBank BLAST-PRODOM BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS ProfileScan
32	326	T50 T84 S98 S142 T261 T65 T148 T178 T189 T221	N187	G-beta WD repeat domain: L120-N153, I140-L154		BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS
33	453	T157 T218 T248 S320 S347 S412 S7 T236 S290 T396 T406 Y63	N59 N225	G-beta WD repeat domain: D180-E211, A198-V212		BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
34	161	T137 T18 T102 Y96		DMR-N9 protein: K93-S148	DMR-N9 (homology to WD repeat sequences) [Mus musculus] g817954	BLAST-GenBank BLAST-PRODOM MOTIFS
35	684	T173 S25 S43 S74 S83 S127 S152 S154 S182 T316 T331 T341 S372 T535 T606 S623 T138 T151 S168 S238 S299 T336 T422 S476 T506 T530 T628 T647	N526 N621	ATP/GTP-binding site motif A (P-loop): G267 Elongation factor 1 alpha protein (GTP-binding) domain: D485-E684 Elongation factor Tu domain: K258-D658, N262-K273, M343-G374, R664-G677 GTP-binding elongation factors signature: A249-E420, N262-T275, K294-P346, T341-F351, T357-V368, L401-Q410, P443-I682 RAS transforming protein: K258-V439	eRFS (related to eukaryotic release factor 3) [Mus musculus] g4566435	BLAST-GenBank BLAST-DOMO BLAST-PRODOM BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS ProfileScan
36	366	S342 T52 S71 T102 T119 T224 T324 T66 S195 S271 T353 Y225	N32	G-beta WD repeat domain: V146-L160, L284-I298 Signal Peptide: M1-T56		BLIMPS-PRINTS MOTIFS SPScan
37	339	S152 S183 T107 T115		Beta-transducin Trp-Asp repeats signature: N101-L162 Trp-Asp repeats-containing protein: R54-A172 Transmembrane domain: A300-I323	Hypothetical trp-asp repeats containing protein [Schizosaccharomyces pombe] g3850059	BLAST-GenBank BLAST-DOMO BLAST-PRODOM HMMER MOTIFS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
38	213	T29 T134 S153 T181 S200 T92 T129 S207		ATP/GTP-binding site motif A (P-loop): G15 GTP-binding protein signature (Arf1, Ran): W5-E179 Ras family signature: R10-C213 Transforming protein p21: F9-E30, R32-R48, E51-S73, Y114-L127, Y149-I171 Signal peptide: M1-V19	Rab-related GTP-binding protein [Homo sapiens] g1491714	BLAST-GenBank BLAST-DOMO BLAST-PRODOM BLIMPS-PRINTS HMMER-PFAM MOTIFS SPScan
39	393	S209 T363 S60 S99 S119 S135 T144 T147 S174 S210 T350 S359 S370 T371		G-beta WD repeat domain: G33-D69, K73-D110, L97-A111, W114-N152, L236-K276, I263-L277 Signal peptide: M1-T43	Similar to beta-transducin [Caenorhabditis elegans] g860695	BLAST-GenBank BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS SPScan
40	399	S86 T191 S219 S224 S254 S275 S308 S59 S72 T96 S373 S385 T394	N88 N106 N321 N322	ATP/GTP-binding site motif A (P-loop): G68 G-protein alpha subunit: R63-Q78 GTP-binding protein GTR1: A57-D294 Ras transforming protein: K61-L203	Gtr2 homolog, novel small GTPase subfamily [Schizosaccharomyces pombe] g3560242	BLAST-GenBank BLAST-DOMO BLAST-PRODOM BLIMPS-PRINTS MOTIFS
41	412	T106 S337 S391 S29 S30 S41 S130 S154 S207 S231 S326 S82 S97 T212 S220	N367	G-beta WD repeat domain: C184-E217, L204-Y218 Signal peptide: M1-G18	Putative transcriptional regulation protein, trp-asp repeat containing [Schizosaccharomyces pombe] g3766375	BLAST-GenBank BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS SPScan

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
42	163	S15 S17 S71 T114 Y49			Arf-like 2 binding protein BART1 [Homo sapiens] g4426962	BLAST-GenBank MOTIFS
43	514	S113 T174 S263 S297 S441 S484 S510 T100 S192 T371 T490 Y255		G-beta WD repeat domain: L204-Q236, L462-G495, M482-P496 Beta-transducin Trp-Asp repeats signature: F216-A261	Similarity to guanine nucleotide binding protein [Caenorhabditis elegans] g3878300	BLAST-GenBank BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS ProfileScan
44	67	T30 S15 Y18		G-protein gamma subunit: E2-L67, M9-R24, K10-P57, D45-G62 Prenyl group binding site (CAAX box): V64	G gamma protein [Mus musculus] g7259257	BLAST-GenBank BLIMPS-BLOCKS BLIMPS-PRINTS HMMER-PFAM MOTIFS
45	315	T148 S162 S209 S244 S252 S45 T48 S132 S140 S158 T214 S244	N79	WD40 domains/G-beta repeats: Q15-N53, G57-N95, G99-D137, P143-D179, G223-D263 WD/G-beta profiles: L71-Q116, T114-V161 WD/G-beta repeat signature: V250-L264	Contains similarity to G beta repeats (PROSITE:PS00670) of the beta-transducin family [Caenorhabditis elegans] g1086900	BLAST-GenBank MOTIFS ProfileScan HMMER-PFAM
46	504	T268 T99 T193 S323 S324 T409 T493 T91 T98 T133 T185 T234 T259 T264 T287 T337 S415 S498	N37 N295	WD40 domains/G-beta repeats: A211-D250, E254-S292, A296-A331, G338-D378, R382-D420 WD/G-beta profiles: T396-I442, T268-A316, C355-F400 WD/G-beta signatures: L407-L421, V279-V293 WD repeat protein-like region: I4-A226	Similar to S. cerevisiae PRP19 protein; simliar to G-beta repeat region of guanine nucleotide binding protein [Caenorhabditis elegans] g727450	BLAST-GenBank BLAST-PRODOM MOTIFS BLIMPS-PRINTS ProfileScan HMMER-PFAM

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
47	522	S84 S315 S510 T20 S50 S57 S74 S116 S122 S128 S161 S185 T274 T300 S339 S345 S357 S367 T373 S459 T474 S136 S143 T174 S200 T300 S315 S356 S385 S420 T492	N226 N355		SAPK (stress activated protein kinase) interacting protein (similar to ras inhibitor) [Gallus gallus] g4929812	BLAST-GenBank MOTIFS
48	316	T109 S27 S86 S188 S7 S8 S82 T96 T105	N29 N136 N186	Pleckstrin homology (PH) domains: S3-N45, I59-Q301 RhoGAP domain: P140-N291 GTPase protein-like region: G125-L307	Beta2-chimaerin [Homo sapiens] g457230	BLAST-GenBank BLAST-PRODOM BLAST-DOMO HMMER-PFAM MOTIFS BLIMPS-PRINTS BLIMPS-PRODOM
49	387	S97 S199 T249 S342 S369 S382 T54 T182 T381		ATP/GTP-binding site motif (P-loop): G155-S162 GTP1/OBG GTP-binding protein family signatures: V151-A171, K172-I190, V200-G215, G217-D235 GTP-binding protein-like region: F15-P173 RAS transforming protein-like region: L145-L296	GTP-binding protein [Aquifex aeolicus] g2984292	BLAST-GenBank BLAST-PRODOM BLAST-DOMO BLIMPS-BLOCKS BLIMPS-PRINTS MOTIFS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
50	334	T228 T308 S65 S91 T224 T228 T262 S34 S81 T224 T262 S286 T324	N108 N257 N322	ATP/GTP-binding site motif (P-loop): G149-S156 Ras domain: R144-M334 p21/ras-related transforming protein signatures: Y143-S164, N166-L182, H248-D261, F282-K304 Ras transforming protein-like region: I140-E284	NOEY2 putative tumor suppressor [Homo sapiens] g4100355	BLAST-GenBank BLAST-PRODOM BLAST-DOMO HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS MOTIFS
51	551	T199 S38 T62 S85 T116 S169 S351 T379 S421 S422 S456 S12 S22 S150 T366 S383 T482 Y404 Y449	N133 N148 N179 N293 N296	Regulator of chromosome condensation (RCC1)/ guanine nucleotide dissociation stimulator domains: E117-S169, D170-D222, T223-D274, E275-G292, G328-G339 RCC1 signatures: V157-L167, V262-L272	UVB-resistance protein UVR8 [Arabidopsis thaliana] g5478530	BLAST-GenBank BLAST-PRODOM HMMER-PFAM PROFILESCAN BLIMPS-PRINTS MOTIFS
52	308	S152 T230 S266 S299 S19 S22 S240	N76	WD40 domains/G-beta repeats: Q33-R73, W79-T119, W126-K181, W188-T230, P241-K276, S11-A50 Sec13 related/WD repeat protein-like region: R73-I177 WD/G-beta profile: G11-A50	Sec13-related protein [Arabidopsis thaliana] g3150415	BLAST-GenBank HMMER-PFAM PROFILESCAN BLIMPS-PRINTS BLAST-PRODOM MOTIFS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
53	949	S206 S514 T22 S216 T226 S273 T315 S663 T745 T908 T155 S232 S258 T350 S359 S472 S609 S776 S837 S913 Y682 Y862	N114	WD40 domains/G-beta repeats: V199-K237, V248-S284, G287-H326 Drosophila lethal(2) giant larvae tumor suppressor protein signature: K221-P244, A353-E377		HMMER-PFAM BLIMPS-PRINTS MOTIFS
54	227	S11 T113 S173 T155 S173	N38	ATP/GTP-binding site motif (P-loop): G37-T44 Ras family domain: K32-C227 p21/ras-related transforming protein signatures: F31-D52, S54-K70, V72-T94, D134-M147, F169-I191 Ras transforming protein-like region: F27-T172	GTP-binding protein [Bos taurus] g162764	BLAST-GenBank HHMER-PFAM BLIMPS-PRINTS BLAST-DOMO BLAST-PRODOM MOTIFS
55	474	T430 S98 S118 S309 S450 S463 T66 S130 T141 S241 S289 S309 S389 S450	N179 N185	WD40 domains/G-beta repeats: D70-Q109, T120-N159, E164-D202 G-beta repeat signature: L146-V160 WD repeat/coronin protein-like region: I208-Q467	Coronin-2 [Mus musculus] g4895039	BLAST-GenBank HMMER-PFAM BLAST-PRODOM BLAST-DOMO MOTIFS
56	547	S16 T77 S85 S90 S112 S114 T132 S160 T166 T225 S248 S438 S491 S526 S125 S267 T299 T305 S504	N101 N110 N147 N297	WD40 domains/G-beta repeats: G159-N197, C312-A353, G357-D396 WD40/G-beta signatures: V245-A259, L428-T442	Guanine nucleotide-binding protein beta 5 [Mesocricetus auratus] g1001939	BLAST-GenBank HMMER-PFAM BLIMPS-PRINTS MOTIFS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
57	686	T331 S431 T637 S34 S169 S554 S28 S124 S192 S273 S341 T366 S426 S449 S470	N26 N44 N271 N424 N628	G-beta profile: S106-S152	Beta-transducin-like protein [Podospora anserina] g607003	BLAST-GenBank PROFILESCAN HMMER-PFAM
58	93	S15 T2 S3 T24			HP protein (RhoGAP ortholog) [Homo sapiens] g2559002	BLAST-GenBank MOTIFS
59	521	S63 S223 T64 T117 S147 S159 S195 S200 T214 S271 S401 S448 T49 S110 S195 T235 T280 T439	N71 N108 N381	Amino acyl tRNA ligase motif: P173-T183	GTPase activating protein [Schizosaccharomyces pombe] g3150248	BLAST-GenBank MOTIFS
60	751	T287 S543 T61 S275 S345 T430 T474 T565 T676 S705 S726 T727 S57 T63 T70 T287 S345 T389 T432 S458 T479 T518 T538	N344 N640	GTP binding elongation factor Tu family domain: E44-T530 Elongation factor G C-terminus domain: L556-T727 GTP binding elongation factor signatures: N48-T61, Q97-A105, N117-F127, R133-V144, F169-R178	Elongation factor G [Rattus norvegicus] g310102	BLAST-GenBank HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS ProfileScan BLAST-PRODOM BLAST-DOMO MOTIFS
61	666	T492 S615 S619 T35 S142 T177 T212 S224 S270 T353 S403 T456 T471 T500 T550 S560 S572 T378 S403 S496 T509 T608 T611 T625	N75 N582		Rho target rhophilin [Mus musculus] g1176422	BLAST-GenBank MOTIFS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
62	746	S22 T98 S571 T46 S53 S61 S66 S70 S71 T97 S14 S126 S127 T165 T184 T190 S249 S279 S323 S430 S519 S680 S736 S115 T190 T237 S349 S436 T444 S567 S598 S601 T613 S652 T741		WD40 domains/G-beta repeats: T403-E441, R570-H606, Q610-D648, T653-H691, L704-T746, C418-A461 G-beta repeat signature: L428-V442 Trp-Asp repeat protein-like region: S22-L407	Bop1 growth control protein [Mus musculus] g1679772	BLAST-GenBank BLAST-PRODOM BLAST-DOMO MOTIFS BLIMPS-PRINTS ProfileScan HMMER-PFAM
63	212	S105 S142 S148 S162 S167 S44 T56 T101 S162 S190	N131	ATP/GTP-binding site motif (P-loop): G25-T32 Ras family domain: K20-C212 ADP-ribosylation factor family domain: P6-R183 p21/ras-related transforming protein signatures: F19-T40, A42-K58, L60-T82, S122-L135, A158-L180 Ras transforming protein-like region: Y15-I155	Rab19 [Mus musculus] g2598565	BLAST-GenBank HMMER-PFAM BLIMPS-BLOCKS BLIMPS-PRINTS BLAST-DOMO BLAST-PRODOM MOTIFS

Table 2 (cont.)

SEQ ID NO:	Amino Acid Residues	Potential Phosphorylation Sites	Potential Glycosylation Sites	Signature Sequences, Motifs, and Domains	Homologous Sequences	Analytical Methods & Databases
64	307	T275 S276 T15 S25 T99 S164 S201 S6 S270 T293	N196 N291	WD40 domains/G-beta repeats: M1-I49, L60-D98, E102-Q140 Sterile alpha motif (SAM): E161-R225 WD/G-beta signatures: L36-V50, L127-F141 G-beta profile: L74-P122	Hypothetical trp-asp repeats protein [C. elegans] SwissProt Q93847	BLAST-SwissProt HMMER-PFAM BLIMPS-PRINTS ProfileScan MOTIFS
65	378	S137 T167 T193 S202 S237 S276 S290 S310 S362 S82 T150 T158 T199 S362 T368		WD40 domains/G-beta repeats: H72-L110, L116-D155, L241-D279 G-beta profiles: S137-C175, S87-C133, I255-S312	WD repeat protein [Schizosaccharomyces pombe] g5701965	BLAST-GenBank HMMER-PFAM ProfileScan MOTIFS
66	466	S6 T24 S69 T209 S246 S357 T450 S181 S236 S242 T322 T407 T450	N448	RasGEF domain: V197-E397 Guanine nucleotide releasing protein-like region: P201-S432	Putative guanine-nucleotide releasing factor [Drosophila affinis] g2981229	BLAST-GenBank HMMER-PFAM BLAST-PRODOM BLAST-DOMO

Table 3

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
67	434-478	Cardiovascular (0.238) Reproductive (0.238) Hematopoietic/Immune (0.190)	Cancer (0.429) Inflammation/Trauma (0.524) Cell Proliferation (0.095)	pINCY
68	380-424 551-595	Nervous (0.185) Reproductive (0.167) Gastrointestinal (0.148)	Cancer (0.444) Cell Proliferation (0.315) Inflammation/Trauma (0.278)	pINCY
69	433-477	Reproductive (0.429) Nervous (0.142) Hematopoietic/Immune (0.142)	Cancer (0.714) Inflammation/Trauma (0.142)	pINCY
70	684-728	Reproductive (0.333) Nervous (0.178) Cardiovascular (0.111)	Cancer (0.467) Cell Proliferation (0.244) Inflammation/Trauma (0.267)	pINCY
71	219-263	Hematopoietic/Immune (0.257) Reproductive (0.229) Gastrointestinal (0.143)	Cell Proliferation (0.400) Inflammation/Trauma (0.429) Cancer (0.314)	pINCY
72	865-912	Gastrointestinal (0.286) Reproductive (0.286) Cardiovascular (0.238)	Cancer (0.667) Cell Proliferation (0.143) Inflammation/Trauma (0.238)	pINCY
73	900-944	Reproductive (0.229) Hematopoietic/Immune (0.157) Nervous (0.157)	Cancer (0.422) Inflammation/Trauma (0.349) Cell Proliferation (0.205)	pINCY
74	109-153 919-963	Reproductive (0.270) Gastrointestinal (0.162) Cardiovascular (0.135)	Cancer (0.405) Cell Proliferation (0.270) Inflammation/Trauma (0.324)	pINCY
75	1352-1396 1568-1612	Reproductive (0.296) Gastrointestinal (0.167) Nervous (0.167)	Cancer (0.509) Inflammation/Trauma (0.269) Cell Proliferation (0.157)	pINCY
76	541-585 1189-1233	Reproductive (0.238) Cardiovascular (0.190) Gastrointestinal (0.190)	Cancer (0.524) Inflammation/Trauma (0.310) Cell Proliferation (0.143)	PBLUESCRIPT
77	110-154	Reproductive (0.250) Nervous (0.224) Hematopoietic/Immune (0.132) Gastrointestinal (0.132)	Cancer (0.355) Inflammation/Trauma (0.342) Cell Proliferation (0.211)	PSPORT1
78	218-262	Reproductive (0.375) Nervous (0.188) Urologic (0.188)	Cancer (0.562) Inflammation/Trauma (0.250)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
79	380-424	Hematopoietic/Immune (0.227) Nervous (0.227) Reproductive (0.227)	Inflammation/Trauma (0.636) Cancer (0.364)	PSPORT1
80	217-261	Reproductive (0.275) Gastrointestinal (0.196) Nervous (0.196)	Cancer (0.431) Inflammation/Trauma (0.451) Cell Proliferation (0.196)	PSPORT1
81	488-532 812-856	Reproductive (0.301) Nervous (0.151) Gastrointestinal (0.130)	Cancer (0.466) Inflammation/Trauma (0.288) Cell Proliferation (0.151)	pINCY
82	595-639	Reproductive (0.333) Developmental (0.148) Gastrointestinal (0.148)	Cancer (0.444) Cell Proliferation (0.370) Inflammation/Trauma (0.333)	pINCY
83	219-263	Hematopoietic/Immune (0.400) Gastrointestinal (0.200) Cardiovascular (0.100)	Inflammation/Trauma (0.429) Cell Proliferation (0.357) Cancer (0.286)	pINCY
84	164-208	Cardiovascular (0.667) Nervous (0.222) Hematopoietic/Immune (0.111)	Cancer (0.556) Cell Proliferation (0.111)	PBLUESCRIPT
85	487-531 757-801	Reproductive (0.182) Cardiovascular (0.091)	Cancer (0.308) Cell Proliferation (0.231) Inflammation/Trauma (0.154)	pINCY
86	325-369 811-855	Hematopoietic/Immune (0.288) Reproductive (0.197) Cardiovascular (0.136)	Inflammation (0.394) Cancer (0.318) Cell Proliferation (0.212)	pINCY
87	163-207	Reproductive (0.218) Nervous (0.172) Gastrointestinal (0.138)	Cancer (0.448) Cell Proliferation (0.218) Inflammation (0.207)	pINCY
88	362-406 758-802	Reproductive (0.273) Gastrointestinal (0.227) Cardiovascular (0.136) Musculoskeletal (0.136)	Cancer (0.681) Cell Proliferation (0.182) Inflammation/Trauma (0.318)	pINCY
89	272-316	Reproductive (0.229) Gastrointestinal (0.193) Nervous (0.193)	Cancer (0.404) Inflammation (0.220) Cell Proliferation (0.165)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
90	98-142	Nervous (0.400) Cardiovascular (0.200) Developmental (0.200) Gastrointestinal (0.200)	Cell Proliferation (0.400) Inflammation (0.400) Cancer (0.200)	pINCY
91	384-428 2016-2060	Reproductive (0.221) Gastrointestinal (0.156) Hematopoietic/Immune (0.143)	Cancer (0.468) Inflammation/Trauma (0.325) Cell Proliferation (0.273)	PBLUESCRIPT
92	80-124 731-775	Reproductive (0.286) Hematopoietic/Immune (0.143) Nervous (0.143)	Cancer (0.469) Inflammation/Trauma (0.326) Cell Proliferation (0.306)	PBLUESCRIPT
93	437-481 641-685	Reproductive (0.250) Nervous (0.200) Cardiovascular (0.183)	Cancer (0.550) Inflammation/Trauma (0.284) Cell Proliferation (0.150)	PBLUESCRIPT
94	397-441 1036-1080	Reproductive (0.291) Hematopoietic/Immune (0.228) Nervous (0.152)	Inflammation/Trauma (0.468) Cancer (0.392) Cell Proliferation (0.165)	pINCY
95	247-291	Reproductive (0.242) Hematopoietic/Immune (0.121) Nervous (0.121) Urologic (0.121)	Cancer (0.455) Inflammation/Trauma (0.333) Cell Proliferation (0.273)	pINCY
96	453-497 858-902	Nervous (0.600) Reproductive (0.400)	Cancer (0.400) Inflammation/Trauma (0.200) Neurological (0.200)	pINCY
97	224-268 770-814 1211-1255	Gastrointestinal (0.262) Reproductive (0.215) Nervous (0.169)	Cancer (0.462) Inflammation/Trauma (0.339) Cell Proliferation (0.231)	pINCY
98	3-47 1086-1130	Reproductive (0.211) Gastrointestinal (0.211) Hematopoietic/Immune (0.158)	Cancer (0.553) Cell Proliferation (0.368) Inflammation/Trauma (0.342)	pINCY
99	388-432 874-918	Reproductive (0.268) Nervous (0.146) Cardiovascular (0.146)	Cancer (0.390) Inflammation/Trauma (0.390) Cell Proliferation (0.220)	pINCY
100	26-70	Gastrointestinal (0.238) Cardiovascular (0.190) Hematopoietic/Immune (0.143) Nervous (0.143) Endocrine (0.143)	Cancer (0.429) Inflammation/Trauma (0.381) Cell Proliferation (0.190)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
101	226-270 2062-2106	Nervous (0.234) Hematopoietic/Immune (0.170) Reproductive (0.149)	Inflammation/Trauma (0.383) Cancer (0.362) Cell Proliferation (0.213)	pINCY
102	487-531	Reproductive (0.276) Nervous (0.161) Gastrointestinal (0.138) Cardiovascular (0.138)	Cancer (0.494) Cell Proliferation (0.310) Inflammation/Trauma (0.264)	pINCY
103	561-605	Reproductive (0.274) Gastrointestinal (0.194) Cardiovascular (0.129)	Cancer (0.452) Inflammation/Trauma (0.339) Cell Proliferation (0.258)	pINCY
104	287-331 806-850	Gastrointestinal (0.500) Reproductive (0.250) Musculoskeletal (0.250)	Cancer (0.500) Inflammation/Trauma (0.250)	pINCY
105	154-198 505-549 757-801	Gastrointestinal (0.233) Reproductive (0.209) Hematopoietic/Immune (0.163) Nervous (0.163)	Cancer (0.465) Inflammation/Trauma (0.326) Cell Proliferation (0.209)	pINCY
106	174-218 1182-1226	Reproductive (0.185) Hematopoietic/Immune (0.185) Nervous (0.185)	Inflammation/Trauma (0.352) Cell Proliferation (0.333) Cancer (0.315)	pINCY
107	120-164 489-533	Reproductive (0.231) Hematopoietic/Immune (0.231) Nervous (0.154) Cardiovascular (0.154)	Cell Proliferation (0.462) Inflammation/Trauma (0.385) Cancer (0.231)	pINCY
108	64-108 1738-1782	Nervous (0.277) Reproductive (0.255) Cardiovascular (0.160)	Cancer (0.362) Inflammation/Trauma (0.362) Cell Proliferation (0.149)	pINCY
109	415-459 1027-1071 1549-1593	Reproductive (0.274) Hematopoietic/Immune (0.226) Nervous (0.167)	Inflammation/Trauma (0.476) Cancer (0.393) Cell Proliferation (0.179)	pINCY
110	242-286	Reproductive (0.500) Nervous (0.500)	Cancer (1.000)	pINCY
111	488-541 1028-1081	Reproductive (0.270) Nervous (0.191) Gastrointestinal (0.126)	Cancer (0.507) Inflammation/Trauma (0.284) Cell Proliferation (0.172)	PSPORT1

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
112	273-326 867-920 1299-1352	Reproductive (0.312) Nervous (0.281) Gastrointestinal (0.094)	Cancer (0.469) Inflammation/Trauma (0.328) Cell Proliferation (0.172)	pINCY
113	866-1135	Reproductive (0.245) Gastrointestinal (0.136) Nervous (0.136)	Cancer (0.445) Cell Proliferation (0.227) Inflammation/Trauma (0.327)	pINCY
114	155-325 812-1105	Nervous (0.314) Reproductive (0.275) Gastrointestinal (0.098)	Cancer (0.471) Inflammation/Trauma (0.118)	pINCY
115	14-298	Gastrointestinal (0.190) Nervous (0.190) Reproductive (0.190)	Cancer (0.476) Cell Proliferation (0.190) Inflammation/Trauma (0.238)	pINCY
116	41-235	Reproductive (0.400) Nervous (0.267) Musculoskeletal (0.133)	Cancer (0.600) Inflammation/Trauma (0.334) Cell Proliferation (0.067)	PSPORT1
117	379-432 973-1026 1297-1350	Reproductive (0.327) Nervous (0.184) Urologic (0.102)	Cancer (0.531) Cell Proliferation (0.224) Inflammation/Trauma (0.265)	pINCY
118	974-1465	Reproductive (0.231) Nervous (0.190) Gastrointestinal (0.169)	Cancer (0.446) Inflammation/Trauma (0.343) Cell Proliferation (0.226)	pINCY
119	543-1028	Reproductive (0.292) Nervous (0.163) Gastrointestinal (0.139)	Cancer (0.517) Cell Proliferation (0.167) Inflammation/Trauma (0.235)	PSPORT1
120	385-552	Nervous (0.571) Cardiovascular (0.143) Developmental (0.143)	Cancer (0.429) Inflammation/Trauma (0.572) Cell Proliferation (0.143)	pINCY
121	685-864	Nervous (0.300) Hematopoietic/Immune (0.200) Cardiovascular (0.140)	Cancer (0.340) Inflammation/Trauma (0.440) Cell Proliferation (0.200)	pINCY
122	703-1026	Reproductive (0.400) Cardiovascular (0.160) Nervous (0.160)	Cancer (0.680) Cell Proliferation (0.120) Inflammation/Trauma (0.160)	pINCY
123	830-1351	Reproductive (0.200) Cardiovascular (0.154) Hematopoietic/Immune (0.154)	Cancer (0.415) Cell Proliferation (0.277) Inflammation/Trauma (0.354)	pINCY

Table 3 (cont.)

Nucleotide SEQ ID NO:	Selected Fragments	Tissue Expression (Fraction of Total)	Disease or Condition (Fraction of Total)	Vector
124	272-325	Cardiovascular (0.250) Gastrointestinal (0.250) Musculoskeletal (0.250)	Inflammation/Trauma (0.750)	pINCY
125	130-972	Reproductive (0.180) Cardiovascular (0.160) Hematopoietic/Immune (0.160)	Cancer (0.440) Inflammation/Trauma (0.340) Cell Proliferation (0.220)	pINCY
126	434-973	Reproductive (0.188) Cardiovascular (0.156) Gastrointestinal (0.156)	Cancer (0.422) Inflammation/Trauma (0.328) Cell Proliferation (0.203)	pINCY
127	489-899	Gastrointestinal (0.333) Reproductive (0.333) Nervous (0.125)	Cancer (0.625) Inflammation/Trauma (0.208) Cell Proliferation (0.042)	pINCY
128	19-1242	Reproductive (0.354) Nervous (0.188) Gastrointestinal (0.146)	Cancer (0.562) Cell Proliferation (0.250) Inflammation/Trauma (0.250)	pINCY
129	217-270 541-594	Reproductive (0.364) Cardiovascular (0.182) Gastrointestinal (0.182)	Cancer (0.636) Inflammation/Trauma (0.364)	pINCY
130	115-864	Gastrointestinal (0.250) Hematopoietic/Immune (0.208) Nervous (0.208)	Cancer (0.500) Inflammation/Trauma (0.292)	pINCY
131	255-308	Reproductive (0.265) Nervous (0.169) Gastrointestinal (0.120)	Cancer (0.482) Cell Proliferation (0.349) Inflammation/Trauma (0.253)	pINCY
132	23-541	Nervous (0.909) Endocrine (0.091)	Cancer (0.636) Cell Proliferation (0.091) Inflammation/Trauma (0.182)	pINCY

Table 4

SEQ ID NO:	Library	Library Comment
67	LATRTUT02	Library was constructed using RNA isolated from a myxoma removed from the left atrium of a 43-year-old Caucasian male during annuloplasty. Pathology indicated atrial myxoma. Patient history included pulmonary insufficiency, acute myocardial infarction, atherosclerotic coronary artery disease, and hyperlipidemia. Family history included benign hypertension, acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
68	PENITUT01	Library was constructed using RNA isolated from tumor tissue removed from the penis of a 64-year-old Caucasian male during penile amputation. Pathology indicated a fungating invasive grade 4 squamous cell carcinoma involving the inner wall of the foreskin and extending onto the glans penis. Patient history included benign neoplasm of the large bowel, atherosclerotic coronary artery disease, angina pectoris, gout, and obesity. Family history included malignant pharyngeal neoplasm, chronic lymphocytic leukemia, and chronic liver disease.
69	BLADTUT04	Library was constructed using RNA isolated from bladder tumor tissue removed from a 60-year-old Caucasian male during a radical cystectomy, prostatectomy, and vasectomy. Pathology indicated grade 3 transitional cell carcinoma in the left bladder wall. Carcinoma in-situ was identified in the dome and trigone. Family history included type I diabetes, a malignant neoplasm of the stomach, atherosclerotic coronary artery disease, and an acute myocardial infarction.
70	BLADTUT06	Library was constructed using RNA isolated from bladder tumor tissue removed from the posterior bladder wall of a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, and gastrostomy. Pathology indicated grade 3 transitional cell carcinoma in the left lateral bladder wall. The remaining bladder showed marked cystitis with scattered microscopic foci of transitional cell carcinoma in situ. Patient history included angina and emphysema. Family history included acute myocardial infarction, atherosclerotic coronary artery disease, and type II diabetes.
71	ADRENOT07	Library was constructed using RNA isolated from adrenal tissue removed from a 61-year-old female during a bilateral adrenalectomy. Patient history included an unspecified disorder of the adrenal glands.
72	BRSTNOT19	Library was constructed using RNA isolated from breast tissue removed from a 67-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology for the associated tumor tissue indicated residual invasive lobular carcinoma. Patient history included depressive disorder, benign large bowel neoplasm, and hemorrhoids. Family history included cerebrovascular and cardiovascular disease and lung cancer.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
73	SPLNNOT12	Library was constructed using RNA isolated from spleen tissue removed from a 65-year-old female. Pathology indicated the spleen was negative for metastasis. Pathology for the associated tumor tissue indicated well-differentiated neuroendocrine carcinoma (islet cell tumor), nuclear grade 1, forming a dominant mass in the distal pancreas. Multiple smaller tumor nodules were immediately adjacent to the main mass. The liver showed metastatic grade 1 islet cell tumor, forming multiple nodules. Multiple (4) pericholedochal lymph nodes contained metastatic grade 1 islet cell tumor.
74	MONOTXT02	Library was constructed using RNA isolated from treated monocytes from peripheral blood removed from a 42-year-old female. The cells were treated with interleukin-10 (IL-10) and lipopolysaccharide (LPS). IL-10 was added at time 0 at 10 ng/ml, LPS was added at 1 hour at 5 ng/ml. The monocytes were isolated from buffy coat by adherence to plastic. Incubation time was 24 hours.
75	FIBPFEN06	Library was constructed from 1.56 million independent clones from a prostate stromal fibroblast tissue library. Starting RNA was made from fibroblasts of prostate stroma removed from a male fetus,who died after 26 weeks' gestation. The libraries were normalized in two rounds using conditions adapted from Soares et al. (1994) Proc. Natl. Acad. Sci. USA 91:9228 and Bonaldo et al.(1996) Genome Research 6:791, except that a significantly longer (48-hours/round) reannealing hybridization was used.
76	HUVESTB01	Library was constructed using RNA isolated from shear-stressed HUV-EC-C (ATCC CRL 1730) cells. Before RNA isolation, the cells were subjected to a shear stress of 10 dynes/cm.
77	SYNOOAT01	Library was constructed using RNA isolated from the knee synovial membrane tissue of an 82-year-old female with osteoarthritis.
78	UTRSNOT05	Library was constructed using RNA isolated from the uterine tissue of a 45-year-old Caucasian female during a total abdominal hysterectomy and total colectomy. Pathology for the associated tumor tissue indicated multiple leiomyomas of the myometrium and a grade 2 colonic adenocarcinoma of the cecum. Patient history included multiple sclerosis and mitral valve disorder. Family history included type I diabetes, cerebrovascular disease, atherosclerotic coronary artery disease, malignant skin neoplasm, hypertension, and malignant neoplasm of the colon.
79	HIPONON01	Library was constructed from 1.13 million independent clones from a hippocampus library. RNA was isolated from the hippocampus tissue of a 72-year-old Caucasian female who died from an intracranial bleed. Patient history included nose cancer, hypertension, and arthritis. The normalization and hybridization conditions were adapted from Soares et al. (1994) Proc. Natl. Acad. Sci. USA 91:9228.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
80	BRSTTUT03	Library was constructed using RNA isolated from breast tumor tissue removed from a 58-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated multicentric invasive grade 4 lobular carcinoma. The mass was identified in the upper outer quadrant, and three separate nodules were found in the lower outer quadrant of the left breast. Patient history included skin cancer, rheumatic heart disease, osteoarthritis, and tuberculosis. Family history included cerebrovascular disease, coronary artery aneurysm, breast cancer, prostate cancer, atherosclerotic coronary artery disease, and type I diabetes.
81	SININOT01	Library was constructed using RNA isolated from ileum tissue obtained from the small intestine of a 4-year-old Caucasian female, who died from a closed head injury. Patient history included jaundice. Previous surgeries included a double hernia repair.
82	SINTFET03	Library was constructed using RNA isolated from small intestine tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation.
83	HNT3AZT01	Library was constructed using RNA isolated from the hNT2 cell line (derived from a human teratocarcinoma that exhibited properties characteristic of a committed neuronal precursor). Cells were treated for three days with 0.35 micromolar 5-aza-2'-deoxycytidine (AZ).
84	ENDANOT01	Library was constructed using RNA isolated from aortic endothelial cell tissue from an explanted heart removed from a male during a heart transplant.
85	LUNGTUT08	Library was constructed using RNA isolated from lung tumor tissue removed from a 63-year-old Caucasian male during a right upper lobectomy with fiberoptic bronchoscopy. Pathology indicated a grade 3 adenocarcinoma. Patient history included atherosclerotic coronary artery disease, an acute myocardial infarction, rectal cancer, an asymtomatic abdominal aortic aneurysm, tobacco abuse, and cardiac dysrhythmia. Family history included congestive heart failure, stomach cancer, and lung cancer, type II diabetes, atherosclerotic coronary artery disease, and an acute myocardial infarction.
86	OVARTUT10	Library was constructed using RNA isolated from ovarian tumor tissue removed from the left ovary of a 58-year-old Caucasian female during a total abdominal hysterectomy, removal of a solitary ovary, and repair of inguinal hernia. Pathology indicated a metastatic grade 3 adenocarcinoma of colonic origin, forming a partially cystic and necrotic tumor mass in the left ovary, and an adenocarcinoma of colonic origin, forming a nodule in the left mesovarium. A single intramural leiomyoma was identified in the myometrium. The cervix showed mild chronic cystic cervicitis. Patient history included benign hypertension, follicular cyst of the ovary, colon cancer, benign colon neoplasm, and osteoarthritis. Family history included emphysema, myocardial infarction, atherosclerotic coronary artery disease, benign hypertension, and hyperlipidemia.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
87	BRSTNOT13	Library was constructed using RNA isolated from breast tissue removed from a 36-year-old Caucasian female during bilateral simple mastectomy. Patient history included a breast neoplasm, depressive disorder, hyperlipidemia, and a chronic stomach ulcer. Family history included cardiovascular and cerebrovascular disease; hyperlipidemia; skin, breast, esophageal, bladder, and bone cancer; and Hodgkin's lymphoma.
88	UTRSNOR01	Library was constructed using RNA isolated from uterine endometrium tissue removed from a 29-year-old Caucasian female during a vaginal hysterectomy and cystocele repair. Pathology indicated the endometrium was secretory, and the cervix showed mild chronic cervicitis with focal squamous metaplasia. Pathology for the associated tumor tissue indicated intramural uterine leiomyoma. Patient history included hypothyroidism, pelvic floor relaxation, and paraplegia. Family history included benign hypertension, type II diabetes, and hyperlipidemia.
89	BRSTTMT02	Library was constructed using RNA isolated from diseased right breast tissue removed from a 46-year-old Caucasian female during a unilateral extended simple mastectomy and open breast biopsy. Pathology indicated mildly proliferative fibrocystic change, including intraductal duct ectasia, papilloma formation, and ductal hyperplasia. Pathology for the associated tumor tissue indicated multifocal ductal carcinoma in situ, both comedo and non-comedo types, nuclear grade 2 with extensive intraductal calcifications. Patient history included deficiency anemia, normal delivery, chronic sinusitis, extrinsic asthma, and kidney infection. Family history included type II diabetes, benign hypertension,cerebrovascular disease, skin cancer, and hyperlipidemia.
90	LIVRDIR01	Library was constructed using RNA isolated from diseased liver tissue removed from a 63-year-old Caucasian female during a liver transplant.Patient history included primary biliary cirrhosis. Serology was positive for anti-mitochondrial antibody.
91	HUVENOB01	Library was constructed using RNA isolated from HUV-EC-C (ATCC CRL 1730) cells.
92	TESTNOT03	Library was constructed using RNA isolated from testicular tissue removed from a 37-year-old Caucasian male, who died from liver disease. Patient history included cirrhosis, jaundice, and liver failure.
93	LUNGNOT02	Library was constructed using RNA isolated from the lung tissue of a 47-year-old Caucasian male, who died of a subarachnoid hemorrhage.
94	LUNGFET03	Library was constructed using RNA isolated from lung tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation.
95	PANCNOT07	Library was constructed using RNA isolated from the pancreatic tissue of a Caucasian male fetus, who died at 23 weeks' gestation.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
96	BRAINOT12	Library was constructed using RNA isolated from brain tissue removed from the right frontal lobe of a 5-year-old Caucasian male during a hemispherectomy. Pathology indicated extensive polymicrogyria and mild to moderate gliosis (predominantly subpial and subcortical), which are consistent with chronic seizure disorder. Family history included a cervical neoplasm.
97	LIVRTUT01	Library was constructed using RNA isolated from liver tumor tissue removed from a 51-year-old Caucasian female during a hepatic lobectomy. Pathology indicated metastatic grade 3 adenocarcinoma consistent with colon cancer. Family history included a malignant neoplasm of the liver.
98	GBLATUT01	Library was constructed using RNA isolated from gall bladder tumor tissue removed from a 78-year-old Caucasian female during a cholecystectomy. Pathology indicated invasive grade 2 squamous cell carcinoma, forming a mass in the gall bladder. Patient history included diverticulitis of the colon, palpitations, benign hypertension, and hyperlipidemia. Family history included a cholecystectomy, atherosclerotic coronary artery disease, hyperlipidemia, and benign hypertension.
99	LEUKNOT02	Library was constructed using RNA isolated from white blood cells of a 45-year-old female with blood type O+. The donor tested positive for cytomegalovirus (CMV).
100	LUNGNOT22	Library was constructed using RNA isolated from lung tissue removed from a 58-year-old Caucasian female. The tissue sample used to construct this library was found to have tumor contaminant upon microscopic examination. Pathology for the associated tumor tissue indicated a caseating granuloma. Family history included congestive heart failure, breast cancer, secondary bone cancer, acute myocardial infarction and atherosclerotic coronary artery disease.
101	ADRETUT06	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 57-year-old Caucasian female during a unilateral right adrenalectomy. Pathology indicated pheochromocytoma, forming a nodular mass completely replacing the medulla of the adrenal gland.
102	ADRETUT06	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 57-year-old Caucasian female during a unilateral right adrenalectomy. Pathology indicated pheochromocytoma, forming a nodular mass completely replacing the medulla of the adrenal gland.
103	THYRNOT10	Library was constructed using RNA isolated from diseased left thyroid tissue removed from a 30-year-old Caucasian female during a unilateral thyroid lobectomy and parathyroid reimplantation. Pathology indicated lymphocytic thyroiditis.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
104	CONNTUT05	Library was constructed using RNA isolated from tumorous skull soft tissue removed from a 34-year-old Caucasian female during skull lesion excision. Pathology indicated grade 3 ependymoma forming an implant in the dermis and subcutis associated with dense fibrosis. Patient history included seizures, bone cancer, and brain cancer. Surgeries included cranioplasty and cerebral meninges lesion excision, and treatment included whole brain radiation. Family history included anxiety and depression.
105	HEAANOT01	Library was constructed using RNA isolated from right coronary and right circumflex coronary artery tissue removed from the explanted heart of a 46-year-old Caucasian male during a heart transplantation. Patient history included myocardial infarction from total occlusion of the left anterior descending coronary artery, atherosclerotic coronary artery disease, hyperlipidemia, myocardial ischemia, dilated cardiomyopathy, left ventricular dysfunction, and tobacco abuse. Family history included atherosclerotic coronary artery disease.
106	UTRMTMT01	Library was constructed using RNA isolated from myometrial tissue removed from a 45-year-old Caucasian female during vaginal hysterectomy and bilateral salpingo-oophorectomy. Pathology indicated the myometrium was negative for tumor. Pathology for the associated tumor tissue indicated multiple (23) subserosal, intramural, and submucosal leiomyomata. The endometrium was in proliferative phase. The right ovary contained an old corpus luteum. The cervix, left ovary, and right and left fallopian tubes were unremarkable. The patient presented with stress incontinence. Patient history included extrinsic asthma without status asthmaticus and normal delivery. Patient medications included Motrin, iron sulfate, Premarin, prednisone, Tylenol #3, and Colace. Family history included cerebrovascular disease, depression, and atherosclerotic coronary artery disease.
107	FIBPFEN06	This normalized library was constructed from 1.56 million independent clones from a prostate stromal fibroblast library. RNA was isolated from a male fetus, who died after 26 weeks' gestation. The normalization and hybridization conditions were adapted from Soares et al. (1994) Proc. Natl. Acad. Sci. USA 91:9228.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
108	BRAINOT19	Library was constructed using RNA isolated from diseased brain tissue removed from the left frontal lobe of a 27-year-old Caucasian male during a brain lobectomy. Pathology indicated a focal deep white matter lesion, characterized by marked gliosis, calcifications, and hemosiderin-laden macrophages, consistent with a remote perinatal injury. This tissue also showed mild to moderate generalized gliosis, predominantly subpial and subcortical, consistent with chronic seizure disorder. The left temporal lobe, including the mesial temporal structures, showed focal, marked pyramidal cell loss and gliosis in hippocampal sector CA1, consistent with mesial temporal sclerosis. GFAP was positive for astrocytes. Patient presented with intractable epilepsy, focal epilepsy, hemiplegia, and an unspecified brain injury. Patient history included cerebral palsy, abnormality of gait, and depressive disorder. Family history included brain cancer.
109	COLCDIT03	Library was constructed using RNA isolated from diseased colon polyp tissue removed from the cecum of a 67-year-old female. Pathology indicated a benign cecum polyp. Pathology for the associated tumor tissue indicated invasive grade 3 adenocarcinoma that arose in tubulovillous adenoma forming a fungating mass in the cecum.
110	BRAXNOT03	Library was constructed using RNA isolated from sensory-motor cortex tissue removed from the brain of a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. The cerebral hemisphere revealed moderate fibrosis of the leptomeninges with focal calcifications. There was evidence of shrunken and slightly eosinophilic pyramidal neurons throughout the cerebral hemispheres. There were also multiple small microscopic areas of cavitation with surrounding gliosis, scattered throughout the cerebral cortex. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly and an enlarged spleen and liver. Patient medications included Simethicone, Lasix, Digoxin, Colace, Zantac, Captopril, and Vasotec.
111	BRAITUT02	Library was constructed using RNA isolated from brain tumor tissue removed from the frontal lobe of a 58-year-old Caucasian male during excision of a cerebral meningeal lesion. Pathology indicated a grade 2 metastatic hypernephroma. Patient history included a grade 2 renal cell carcinoma, insomnia, and chronic airway obstruction. Family history included a malignant neoplasm of the kidney.
112	PROSNOT11	Library was constructed using RNA isolated from the prostate tissue of a 28-year-old Caucasian male, who died from a self-inflicted gunshot wound.
113	LIVRTUT01	Library was constructed using RNA isolated from liver tumor tissue removed from a 51-year-old Caucasian female during a hepatic lobectomy. Pathology indicated metastatic grade 3 adenocarcinoma consistent with colon cancer. Family history included a malignant neoplasm of the liver.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
114	PANCTUT02	Library was constructed using RNA isolated from pancreatic tumor tissue removed from a 45-year-old Caucasian female during radical pancreaticoduodenectomy. Pathology indicated a grade 4 anaplastic carcinoma. Family history included benign hypertension, hyperlipidemia and atherosclerotic coronary artery disease.
115	LIVRFET02	Library was constructed using RNA isolated from liver tissue removed from a Caucasian female fetus, who died at 20 weeks' gestation. Family history included seven days of erythromycin treatment for bronchitis in the mother during the first trimester.
116	BRAITUT03	Library was constructed using RNA isolated from brain tumor tissue removed from the left frontal lobe a 17-year-old Caucasian female during excision of a cerebral meningeal lesion. Pathology indicated a grade 4 fibrillary giant and small-cell astrocytoma. Family history included benign hypertension and cerebrovascular disease.
117	BRSTNOT07	Library was constructed using RNA isolated from diseased breast tissue removed from a 43-year-old Caucasian female during a unilateral extended simple mastectomy. Pathology indicated mildly proliferative fibrocystic changes with epithelial hyperplasia, papillomatosis, and duct ectasia. Pathology for the associated tumor tissue indicated invasive grade 4, nuclear grade 3 mammary adenocarcinoma with extensive comedo necrosis. Family history included epilepsy, cardiovascular disease, and type II diabetes.
118	SMCANOT01	Library was constructed using RNA isolated from an aortic smooth muscle cell line derived from the explanted heart of a male during a heart transplant.
119	THP1AZS08	Library was constructed using 5.76 million clones from a 5-aza-2'-deoxycytidine (AZ) treated THP-1 promonocyte cell line library. Starting RNA was made from THP-1 promonocyte cells treated for three days with 0.8 micromolar AZ. The hybridization probe for subtraction was derived from a similarly constructed library, made from 1 microgram of polyA RNA isolated from untreated THP-1 cells. 5.76 million clones from the AZ-treated THP-1 cell library were then subjected to two rounds of subtractive hybridization with 5 million clones from the untreated THP-1 cell library. Subtractive hybridization conditions were based on the methodologies of Swaroop et al. (1991) Nucleic Acids Res. 19:1954, and Bonaldo et al. (1996) Genome Research 6:791. THP-1 (ATCC TIB 202) is a human promonocyte line derived from peripheral blood of a 1-year-old Caucasian male with acute monocytic leukemia (ref: Int. J. Cancer (1980) 26:171).
120	ADRETUT06	Library was constructed using RNA isolated from adrenal tumor tissue removed from a 57-year-old Caucasian female during a unilateral right adrenalectomy. Pathology indicated pheochromocytoma, forming a nodular mass completely replacing the medulla of the adrenal gland.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
121	SININOT03	Library was constructed using RNA isolated from ileum tissue obtained from an 8-year-old Caucasian female, who died from head trauma. Serology was positive for cytomegalovirus (CMV).
122	SININOT03	Library was constructed using RNA isolated from ileum tissue obtained from an 8-year-old Caucasian female, who died from head trauma. Serology was positive for cytomegalovirus (CMV).
123	TLYMNOT06	Library was constructed using RNA isolated from activated Th2 cells. These cells were differentiated from umbilical cord CD4 T cells with IL-4 in the presence of anti-IL-12 antibodies and B7-transfected COS cells, and then activated for six hours with anti-CD3 and anti-CD28 antibodies.
124	HEAANOT01	Library was constructed using RNA isolated from right coronary and right circumflex coronary artery tissue removed from the explanted heart of a 46-year-old Caucasian male during a heart transplantation. Patient history included myocardial infarction from total occlusion of the left anterior descending coronary artery, atherosclerotic coronary artery disease, hyperlipidemia, myocardial ischemia, dilated cardiomyopathy, left ventricular dysfunction, and tobacco abuse. Previous surgeries included cardiac catheterization. Family history included atherosclerotic coronary artery disease.
125	TLYJINT01	Library was constructed using RNA isolated from a Jurkat cell line derived from the T cells of a male. The cells were treated for 18 hours with 50 ng/ml phorbol ester (PMA) and 1 micromolar calcium ionophore. Patient history included acute T-cell leukemia.
126	BRAITUT24	Library was constructed using RNA isolated from right frontal brain tumor tissue removed from a 50-year-old Caucasian male during a cerebral meninges lesion excision. Pathology indicated meningioma. Family history included colon cancer and cerebrovascular disease.
127	PROSTUT16	Library was constructed using RNA isolated from prostate tumor tissue removed from a 55-year-old Caucasian male. Pathology indicated adenocarcinoma, Gleason grade 5+4. Adenofibromatous hyperplasia was also present. The patient presented with elevated prostate specific antigen (PSA). Patient history included calculus of the kidney. Family history included lung cancer and breast cancer.
128	BRONNOT01	Library was constructed using RNA isolated from bronchial tissue removed from a 15-year-old Caucasian male.

Table 4 (cont.)

SEQ ID NO:	Library	Library Comment
129	BLADTUT03	Library was constructed using RNA isolated from bladder tumor tissue removed from a 58-year-old Caucasian male during a radical cystectomy, radical prostatectomy, regional lymph node excision, and urinary diversion to bowel. Pathology indicated invasive grade 3 transitional cell carcinoma. Patient history included a benign colon neoplasm. Family history included cerebrovascular disease and atherosclerotic coronary artery disease.
130	COLXTDT01	Library was constructed using RNA isolated from colon tissue removed from the appendix of a 37-year-old Black female during myomectomy, dilation and curettage, right fimbrial region biopsy, and incidental appendectomy. Pathology indicated an unremarkable appendix. Pathology for the associated tumor tissue indicated multiple (12) uterine leiomyomata. Patient history included premenopausal menorrhagia and sarcoidosis of the lung. Family history included acute myocardial infarction and atherosclerotic coronary artery disease.
131	BRATNOT02	Library was constructed using RNA isolated from superior temporal cortex tissue removed from the brain of a 35-year-old Caucasian male. No neuropathology was found. Patient history included dilated cardiomyopathy, congestive heart failure, and an enlarged spleen and liver.
132	BRAWNOT01	Library was constructed using RNA isolated from dentate nucleus tissue removed from the brain of a 35-year-old Caucasian male who died from cardiac failure. Pathology indicated moderate leptomeningeal fibrosis and multiple microinfarctions of the cerebral neocortex. Patient history included dilated cardiomyopathy, congestive heart failure, cardiomegaly, and an enlarged spleen and liver.

Table 5 (cont.)

Program	Description	Reference	Parameter Threshold
ProfileScan	An algorithm that searches for structural and sequence motifs in protein sequences that match sequence patterns defined in Prosite.	Gribskov, M. et al. (1988) CABIOS 4:61-66; Gribskov, M. et al. (1989) Methods Enzymol. 183:146-159; Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221.	Normalized quality score≥GCG-specified "HIGH" value for that particular Prosite motif. Generally, score=1.4-2.1.
Phred	A base-calling algorithm that examines automated sequencer traces with high sensitivity and probability.	Ewing, B. et al. (1998) Genome Res. 8:175-185; Ewing, B. and P. Green (1998) Genome Res. 8:186-194.	
Phrap	A Phils Revised Assembly Program including SWAT and CrossMatch, programs based on efficient implementation of the Smith-Waterman algorithm, useful in searching sequence homology and assembling DNA sequences.	Smith, T.F. and M.S. Waterman (1981) Adv. Appl. Math. 2:482-489; Smith, T.F. and M.S. Waterman (1981) J. Mol. Biol. 147:195-197; and Green, P., University of Washington, Seattle, WA.	Score= 120 or greater; Match length= 56 or greater
Consed	A graphical tool for viewing and editing Phrap assemblies.	Gordon, D. et al. (1998) Genome Res. 8:195-202.	
SPScan	A weight matrix analysis program that scans protein sequences for the presence of secretory signal peptides.	Nielson, H. et al. (1997) Protein Engineering 10:1-6; Claverie, J.M. and S. Audic (1997) CABIOS 12:431-439.	Score=3.5 or greater
Motifs	A program that searches amino acid sequences for patterns that matched those defined in Prosite.	Bairoch, A. et al. (1997) Nucleic Acids Res. 25:217-221; Wisconsin Package Program Manual, version 9, page M51-59, Genetics Computer Group, Madison, WI.	

What is claimed is:

1. An isolated polypeptide comprising an amino acid sequence selected from the group consisting of:

a) an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, and SEQ ID NO:66,

b) a naturally occurring amino acid sequence having at least 90% sequence identity to an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NQ:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, and SEQ ID NO:66,

c) a biologically active fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:52, SEQ ID

NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, and SEQ ID NO:66, and

d) an immunogenic fragment of an amino acid sequence selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, and SEQ ID NO:66.

2. An isolated polypeptide of claim 1 selected from the group consisting of SEQ ID NO:1, SEQ ID NO:2, SEQ ID NO:3, SEQ ID NO:4, SEQ ID NO:5, SEQ ID NO:6, SEQ ID NO:7, SEQ ID NO:8, SEQ ID NO:9, SEQ ID NO:10, SEQ ID NO:11, SEQ ID NO:12, SEQ ID NO:13, SEQ ID NO:14, SEQ ID NO:15, SEQ ID NO:16, SEQ ID NO:17, SEQ ID NO:18, SEQ ID NO:19, SEQ ID NO:20, SEQ ID NO:21, SEQ ID NO:22, SEQ ID NO:24, SEQ ID NO:25, SEQ ID NO:26, SEQ ID NO:27, SEQ ID NO:29, SEQ ID NO:30, SEQ ID NO:31, SEQ ID NO:32, SEQ ID NO:33, SEQ ID NO:34, SEQ ID NO:36, SEQ ID NO:37, SEQ ID NO:38, SEQ ID NO:39, SEQ ID NO:40, SEQ ID NO:41, SEQ ID NO:43, SEQ ID NO:44, SEQ ID NO:45, SEQ ID NO:46, SEQ ID NO:47, SEQ ID NO:48, SEQ ID NO:49, SEQ ID NO:50, SEQ ID NO:52, SEQ ID NO:53, SEQ ID NO:54, SEQ ID NO:55, SEQ ID NO:56, SEQ ID NO:57, SEQ ID NO:58, SEQ ID NO:59, SEQ ID NO:60, SEQ ID NO:61, SEQ ID NO:62, SEQ ID NO:63, SEQ ID NO:64, SEQ ID NO:65, and SEQ ID NO:66.

3. An isolated polynucleotide encoding a polypeptide of claim 1.

4. An isolated polynucleotide encoding a polypeptide of claim 2.

5. An isolated polynucleotide of claim 4 selected from the group consisting of SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID

NO:85, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:110, SEQ ID NO:111, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:119, SEQ ID NO:120, SEQ ID NO:121, SEQ ID NO:122, SEQ ID NO:123, SEQ ID NO:124, SEQ ID NO:125, SEQ ID NO:126, SEQ ID NO:127, SEQ ID NO:128, SEQ ID NO:129, SEQ ID NO:130, SEQ ID NO:131, and SEQ ID NO:132.

6. A recombinant polynucleotide comprising a promoter sequence operably linked to a polynucleotide of claim 3.

7. A cell transformed with a recombinant polynucleotide of claim 6.

8. A transgenic organism comprising a recombinant polynucleotide of claim 6.

9. A method for producing a polypeptide of claim 1, the method comprising:

a) culturing a cell under conditions suitable for expression of the polypeptide, wherein said cell is transformed with a recombinant polynucleotide, and said recombinant polynucleotide comprises a promoter sequence operably linked to a polynucleotide encoding the polypeptide of claim 1, and

b) recovering the polypeptide so expressed.

10. An isolated antibody which specifically binds to a polypeptide of claim 1.

11. An isolated polynucleotide comprising a polynucleotide sequence selected from the group consisting of:

a) a polynucleotide sequence selected from the group consisting of SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:110, SEQ ID NO:111, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:119,

SEQ ID NO:120, SEQ ID NO:121, SEQ ID NO:122, SEQ ID NO:123, SEQ ID NO:124, SEQ ID NO:125, SEQ ID NO:126, SEQ ID NO:127, SEQ ID NO:128, SEQ ID NO:129, SEQ ID NO:130, SEQ ID NO:131, and SEQ ID NO:132,

b) a naturally occurring polynucleotide sequence having at least 70% sequence identity to a polynucleotide sequence selected from the group consisting of SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:110, SEQ ID NO:111, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:119, SEQ ID NO:120, SEQ ID NO:121, SEQ ID NO:122, SEQ ID NO:123, SEQ ID NO:124, SEQ ID NO:125, SEQ ID NO:126, SEQ ID NO:127, SEQ ID NO:128, SEQ ID NO:129, SEQ ID NO:130, SEQ ID NO:131, and SEQ ID NO:132,

c) a polynucleotide sequence complementary to a),

d) a polynucleotide sequence complementary to b), and

e) an RNA equivalent of a)-d).

12. An isolated polynucleotide comprising at least 60 contiguous nucleotides of a polynucleotide of claim 11.

13. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) hybridizing the sample with a probe comprising at least 20 contiguous nucleotides comprising a sequence complementary to said target polynucleotide in the sample, and which probe specifically hybridizes to said target polynucleotide, under conditions whereby a hybridization complex is formed between said probe and said target polynucleotide or fragments thereof, and

b) detecting the presence or absence of said hybridization complex, and, optionally, if present, the amount thereof.

14. A method of claim 13, wherein the probe comprises at least 60 contiguous nucleotides.

15. A method for detecting a target polynucleotide in a sample, said target polynucleotide having a sequence of a polynucleotide of claim 11, the method comprising:

a) amplifying said target polynucleotide or fragment thereof using polymerase chain reaction amplification, and

b) detecting the presence or absence of said amplified target polynucleotide or fragment thereof, and, optionally, if present, the amount thereof.

16. A composition comprising an effective amount of a polypeptide of claim 1 and a pharmaceutically acceptable excipient.

17. A composition of claim 16, wherein the polypeptide comprises an amino acid sequence selected from the group consisting of SEQ ID NO:67, SEQ ID NO:68, SEQ ID NO:69, SEQ ID NO:70, SEQ ID NO:71, SEQ ID NO:72, SEQ ID NO:73, SEQ ID NO:74, SEQ ID NO:75, SEQ ID NO:76, SEQ ID NO:77, SEQ ID NO:78, SEQ ID NO:79, SEQ ID NO:80, SEQ ID NO:81, SEQ ID NO:82, SEQ ID NO:83, SEQ ID NO:84, SEQ ID NO:85, SEQ ID NO:86, SEQ ID NO:87, SEQ ID NO:88, SEQ ID NO:90, SEQ ID NO:91, SEQ ID NO:92, SEQ ID NO:93, SEQ ID NO:95, SEQ ID NO:96, SEQ ID NO:97, SEQ ID NO:98, SEQ ID NO:99, SEQ ID NO:100, SEQ ID NO:102, SEQ ID NO:103, SEQ ID NO:104, SEQ ID NO:105, SEQ ID NO:106, SEQ ID NO:107, SEQ ID NO:109, SEQ ID NO:110, SEQ ID NO:111, SEQ ID NO:112, SEQ ID NO:113, SEQ ID NO:114, SEQ ID NO:115, SEQ ID NO:116, SEQ ID NO:118, SEQ ID NO:119, SEQ ID NO:120, SEQ ID NO:121, SEQ ID NO:122, SEQ ID NO:123, SEQ ID NO:124, SEQ ID NO:125, SEQ ID NO:126, SEQ ID NO:127, SEQ ID NO:128, SEQ ID NO:129, SEQ ID NO:130, SEQ ID NO:131, and SEQ ID NO:132.

18. A method for treating a disease or condition associated with decreased expression of functional GBAP, comprising administering to a patient in need of such treatment the pharmaceutical composition of claim 16.

19. A method for screening a compound for effectiveness as an agonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting agonist activity in the sample.

20. A composition comprising an agonist compound identified by a method of claim 19 and a pharmaceutically acceptable excipient.

21. A method for treating a disease or condition associated with decreased expression of functional GBAP, comprising administering to a patient in need of such treatment a pharmaceutical

composition of claim 20.

22. A method for screening a compound for effectiveness as an antagonist of a polypeptide of claim 1, the method comprising:

a) exposing a sample comprising a polypeptide of claim 1 to a compound, and

b) detecting antagonist activity in the sample.

23. A composition comprising an antagonist compound identified by a method of claim 22 and a pharmaceutically acceptable excipient.

24. A method for treating a disease or condition associated with overexpression of functional GBAP, comprising administering to a patient in need of such treatment a pharmaceutical composition of claim 23.

25. A method of screening for a compound that specifically binds to the polypeptide of claim 1, said method comprising the steps of:

a) combining the polypeptide of claim 1 with at least one test compound under suitable conditions, and

b) detecting binding of the polypeptide of claim 1 to the test compound, thereby identifying a compound that specifically binds to the polypeptide of claim 1.

26. A method of screening for a compound that modulates the activity of the polypeptide of claim 1, said method comprising:

a) combining the polypeptide of claim 1 with at least one test compound under conditions permissive for the activity of the polypeptide of claim 1,

b) assessing the activity of the polypeptide of claim 1 in the presence of the test compound, and

c) comparing the activity of the polypeptide of claim 1 in the presence of the test compound with the activity of the polypeptide of claim 1 in the absence of the test compound, wherein a change in the activity of the polypeptide of claim 1 in the presence of the test compound is indicative of a compound that modulates the activity of the polypeptide of claim 1.

27. A method for screening a compound for effectiveness in altering expression of a target polynucleotide, wherein said target polynucleotide comprises a sequence of claim 5, the method comprising:

a) exposing a sample comprising the target polynucleotide to a compound, and

b) detecting altered expression of the target polynucleotide.

28. A method for assessing toxicity of a test compound, said method comprising:

a) treating a biological sample containing nucleic acids with the test compound;

b) hybridizing the nucleic acids of the treated biological sample with a probe comprising at least 20 contiguous nucleotides of a polynucleotide of claim 11 under conditions whereby a specific hybridization complex is formed between said probe and a target polynucleotide in the biological sample, said target polynucleotide comprising a polynucleotide sequence of a polynucleotide of claim 11 or fragment thereof;

c) quantifying the amount of hybridization complex; and

d) comparing the amount of hybridization complex in the treated biological sample with the amount of hybridization complex in an untreated biological sample, wherein a difference in the amount of hybridization complex in the treated biological sample is indicative of toxicity of the test compound.

SEQUENCE LISTING

<110> INCYTE GENOMICS, INC.
YUE, Henry
TANG, Y. Tom
BANDMAN, Olga
HILLMAN, Jennifer L.
LAL, Preeti
AU-YOUNG, Janice
REDDY, Roopa
YANG, Junming
BAUGHN, Mariah R.
LU, Dyung Aina M.
AZIMZAI, Yalda
PATTERSON, Chandra

<120> GTP-BINDING ASSOCIATED PROTEINS

<130> PF-0714 PCT

<140> To Be Assigned
<141> Herewith

<150> 60/144,595; 60/150,460; 60/159,849
<151> 1999-07-19; 1999-08-23; 1999-10-15

<160> 132
<170> PERL Program

<210> 1
<211> 269
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1405545CD1

<400> 1
```
Met Pro Ala Val Leu Glu Arg Leu Ser Arg Tyr Asn Ser Thr Ser
  1               5                  10                  15
Gln Ala Phe Ala Glu Val Leu Arg Leu Pro Lys Gln Gln Leu Arg
                 20                  25                  30
Lys Leu Leu Tyr Pro Leu Gln Glu Val Glu Arg Phe Leu Ala Pro
                 35                  40                  45
Tyr Gly Arg Gln Asp Leu His Leu Arg Ile Phe Asp Pro Ser Pro
                 50                  55                  60
Glu Asp Ile Ala Arg Ala Asp Asn Ile Phe Thr Ala Thr Glu Arg
                 65                  70                  75
Asn Arg Ile Asp Tyr Val Ser Ser Ala Val Arg Ile Asp His Ala
                 80                  85                  90
Pro Asp Leu Pro Arg Pro Glu Val Cys Phe Ile Gly Arg Ser Asn
                 95                 100                 105
Val Gly Lys Ser Ser Leu Ile Lys Ala Leu Phe Ser Leu Ala Pro
                110                 115                 120
Glu Val Glu Val Arg Val Ser Lys Lys Pro Gly His Thr Lys Lys
                125                 130                 135
Met Asn Phe Phe Lys Val Gly Lys His Phe Thr Val Val Asp Met
                140                 145                 150
Pro Gly Tyr Gly Phe Arg Ala Pro Glu Asp Phe Val Asp Met Val
                155                 160                 165
Glu Thr Tyr Leu Lys Glu Arg Arg Asn Leu Lys Arg Thr Phe Leu
```

```
                170                 175                 180
Leu Val Asp Ser Val Val Gly Ile Gln Lys Thr Asp Asn Ile Ala
                185                 190                 195
Ile Glu Met Cys Glu Glu Phe Ala Leu Pro Tyr Val Ile Val Leu
                200                 205                 210
Thr Lys Ile Asp Lys Ser Ser Lys Gly His Leu Leu Lys Gln Val
                215                 220                 225
Leu Gln Ile Gln Lys Phe Val Asn Met Lys Thr Gln Gly Cys Phe
                230                 235                 240
Pro Gln Leu Phe Pro Val Ser Ala Val Thr Phe Ser Gly Ile His
                245                 250                 255
Leu Leu Arg Cys Phe Ile Ala Ser Val Thr Gly Ser Leu Asp
                260                 265
<210> 2
<211> 428
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1451265CD1

<400> 2
Met Glu Val Ala Val Cys Thr Asp Ser Ala Ala Pro Met Trp Ser
  1               5                  10                  15
Cys Ile Val Trp Glu Leu His Ser Gly Ala Asn Leu Leu Thr Tyr
                 20                  25                  30
Arg Gly Gly Gln Ala Gly Pro Arg Gly Leu Ala Leu Leu Asn Gly
                 35                  40                  45
Glu Tyr Leu Leu Ala Ala Gln Leu Gly Lys Asn Tyr Ile Ser Ala
                 50                  55                  60
Trp Glu Leu Gln Arg Lys Asp Gln Leu Gln Gln Lys Ile Met Cys
                 65                  70                  75
Pro Gly Pro Val Thr Cys Leu Thr Ala Ser Pro Asn Gly Leu Tyr
                 80                  85                  90
Val Leu Ala Gly Val Ala Glu Ser Ile His Leu Trp Glu Val Ser
                 95                 100                 105
Thr Gly Asn Leu Leu Val Ile Leu Ser Arg His Tyr Gln Asp Val
                110                 115                 120
Ser Cys Leu Gln Phe Thr Gly Asp Ser Ser His Phe Ile Ser Gly
                125                 130                 135
Gly Lys Asp Cys Leu Val Leu Val Trp Ser Leu Cys Ser Val Leu
                140                 145                 150
Gln Ala Asp Pro Ser Arg Ile Pro Ala Pro Arg His Val Trp Ser
                155                 160                 165
His His Thr Leu Pro Ile Thr Asp Leu His Cys Gly Phe Gly Gly
                170                 175                 180
Pro Leu Ala Arg Val Ala Thr Ser Ser Leu Asp Gln Thr Val Lys
                185                 190                 195
Leu Trp Glu Val Ser Ser Gly Glu Leu Leu Leu Ser Val Leu Phe
                200                 205                 210
Asp Val Ser Ile Met Ala Val Thr Met Asp Leu Ala Glu His His
                215                 220                 225
Met Phe Cys Gly Gly Ser Glu Gly Ser Ile Phe Gln Val Asp Leu
                230                 235                 240
Phe Thr Trp Pro Gly Gln Arg Glu Arg Ser Phe His Pro Glu Gln
                245                 250                 255
Asp Ala Gly Lys Val Phe Lys Gly His Arg Asn Gln Val Thr Cys
                260                 265                 270
Leu Ser Val Ser Thr Asp Gly Ser Val Leu Leu Ser Gly Ser His
                275                 280                 285
Asp Glu Thr Val Arg Leu Trp Asp Val Gln Ser Lys Gln Cys Ile
                290                 295                 300
```

Arg Thr Val Ala Leu Lys Gly Pro Val Thr Asn Ala Ala Ile Leu
305 310 315
Leu Ala Pro Val Ser Met Leu Ser Ser Asp Phe Arg Pro Ser Leu
320 325 330
Pro Leu Pro His Phe Asn Lys His Leu Leu Gly Ala Glu His Gly
335 340 345
Asp Glu Pro Arg His Gly Gly Leu Thr Leu Arg Leu Gly Leu His
350 355 360
Gln Gln Gly Ser Glu Pro Ser Tyr Leu Asp Arg Thr Glu Gln Leu
365 370 375
Gln Ala Val Leu Cys Ser Thr Met Glu Lys Ser Val Leu Gly Gly
380 385 390
Gln Asp Gln Leu Arg Val Arg Val Thr Glu Leu Glu Asp Glu Val
395 400 405
Arg Asn Leu Arg Lys Ile Asn Arg Asp Leu Phe Asp Phe Ser Thr
410 415 420
Arg Phe Ile Thr Arg Pro Ala Lys
425

<210> 3
<211> 562
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1556311CD1

<400> 3
Met Pro Glu Thr Val Asn His Asn Lys His Gly Asn Val Ala Leu
1 5 10 15
Pro Gly Thr Lys Pro Thr Pro Ile Pro Pro Pro Arg Leu Lys Lys
20 25 30
Gln Ala Ser Phe Leu Glu Ala Glu Gly Gly Ala Lys Thr Leu Ser
35 40 45
Gly Gly Arg Pro Gly Ala Gly Pro Glu Leu Glu Leu Gly Thr Ala
50 55 60
Gly Ser Pro Gly Gly Ala Pro Pro Glu Ala Ala Pro Gly Asp Cys
65 70 75
Thr Arg Ala Pro Pro Pro Ser Ser Glu Ser Arg Pro Pro Cys His
80 85 90
Gly Gly Arg Gln Arg Leu Ser Asp Met Ser Ile Ser Thr Ser Ser
95 100 105
Ser Asp Ser Leu Glu Phe Asp Arg Ser Met Pro Leu Phe Gly Tyr
110 115 120
Glu Ala Asp Thr Asn Ser Ser Leu Glu Asp Tyr Glu Gly Glu Ser
125 130 135
Asp Gln Glu Thr Met Ala Pro Pro Ile Lys Ser Lys Lys Lys Arg
140 145 150
Ser Ser Ser Phe Val Leu Pro Lys Leu Val Lys Ser Gln Leu Gln
155 160 165
Lys Val Ser Gly Val Phe Ser Ser Phe Met Thr Pro Glu Lys Arg
170 175 180
Met Val Arg Arg Ile Ala Glu Leu Ser Arg Asp Lys Cys Thr Tyr
185 190 195
Phe Gly Cys Leu Val Gln Asp Tyr Val Ser Phe Leu Gln Glu Asn
200 205 210
Lys Glu Cys His Val Ser Ser Thr Asp Met Leu Gln Thr Ile Arg
215 220 225
Gln Phe Met Thr Gln Val Lys Asn Tyr Leu Ser Gln Ser Ser Glu
230 235 240
Leu Asp Pro Pro Ile Glu Ser Leu Ile Pro Glu Asp Gln Ile Asp
245 250 255
Val Val Leu Glu Lys Ala Met His Lys Cys Ile Leu Lys Pro Leu

```
                260                     265                     270
Lys Gly His Val Glu Ala Met Leu Lys Asp Phe His Met Ala Asp
                275                     280                     285
Gly Ser Trp Lys Gln Leu Lys Glu Asn Leu Gln Leu Val Arg Gln
                290                     295                     300
Arg Asn Pro Gln Glu Leu Gly Val Phe Ala Pro Thr Pro Asp Phe
                305                     310                     315
Val Asp Val Glu Lys Ile Lys Val Lys Phe Met Thr Met Gln Lys
                320                     325                     330
Met Tyr Ser Pro Glu Lys Lys Val Met Leu Leu Leu Arg Val Cys
                335                     340                     345
Lys Leu Ile Tyr Thr Val Met Glu Asn Asn Ser Gly Arg Met Tyr
                350                     355                     360
Gly Ala Asp Asp Phe Leu Pro Val Leu Thr Tyr Val Ile Ala Gln
                365                     370                     375
Cys Asp Met Leu Glu Leu Asp Thr Glu Ile Glu Tyr Met Met Glu
                380                     385                     390
Leu Leu Asp Pro Ser Leu Leu His Gly Glu Gly Gly Tyr Tyr Leu
                395                     400                     405
Thr Ser Ala Tyr Gly Ala Leu Ser Leu Ile Lys Asn Phe Gln Glu
                410                     415                     420
Glu Gln Ala Ala Arg Leu Leu Ser Ser Glu Thr Arg Asp Thr Leu
                425                     430                     435
Arg Gln Trp His Lys Arg Arg Thr Thr Asn Arg Thr Ile Pro Ser
                440                     445                     450
Val Asp Asp Phe Gln Asn Tyr Leu Arg Val Ala Phe Gln Glu Val
                455                     460                     465
Asn Ser Gly Cys Thr Gly Lys Thr Leu Leu Val Arg Pro Tyr Ile
                470                     475                     480
Thr Thr Glu Asp Val Cys Gln Ile Cys Ala Glu Lys Phe Lys Val
                485                     490                     495
Gly Asp Pro Glu Glu Tyr Ser Leu Phe Leu Phe Val Asp Glu Thr
                500                     505                     510
Trp Gln Gln Leu Ala Glu Asp Thr Tyr Pro Gln Lys Ile Lys Ala
                515                     520                     525
Glu Leu His Ser Arg Pro Gln Pro His Ile Phe His Phe Val Tyr
                530                     535                     540
Lys Arg Ile Lys Asn Asp Pro Tyr Gly Ile Ile Phe Gln Asn Gly
                545                     550                     555
Glu Glu Asp Leu Thr Thr Ser
                560
```

<210> 4
<211> 229
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1901373CD1

<400> 4

```
Met Ala Glu Asp Lys Thr Lys Pro Ser Glu Leu Asp Gln Gly Lys
  1               5                  10                  15
Tyr Asp Ala Asp Asp Asn Val Lys Ile Ile Cys Leu Gly Asp Ser
                 20                  25                  30
Ala Val Gly Lys Ser Lys Leu Met Glu Arg Phe Leu Met Asp Gly
                 35                  40                  45
Phe Gln Pro Gln Gln Leu Ser Thr Tyr Ala Leu Thr Leu Tyr Lys
                 50                  55                  60
His Thr Ala Thr Val Asp Gly Arg Thr Ile Leu Val Asp Phe Trp
                 65                  70                  75
Asp Thr Ala Gly Gln Glu Arg Phe Gln Ser Met His Ala Ser Tyr
                 80                  85                  90
```

```
Tyr His Lys Ala His Ala Cys Ile Met Val Phe Asp Val Gln Arg
                95                  100                 105
Lys Val Thr Tyr Arg Asn Leu Ser Thr Trp Tyr Thr Glu Leu Arg
                110                 115                 120
Glu Phe Arg Pro Glu Ile Pro Cys Ile Val Val Ala Asn Lys Ile
                125                 130                 135
Asp Ala Asp Ile Asn Val Thr Gln Lys Ser Phe Asn Phe Ala Lys
                140                 145                 150
Lys Phe Ser Leu Pro Leu Tyr Phe Val Ser Ala Ala Asp Gly Thr
                155                 160                 165
Asn Val Val Lys Leu Phe Asn Asp Ala Ile Arg Leu Ala Val Ser
                170                 175                 180
Tyr Lys Gln Asn Ser Gln Asp Phe Met Asp Glu Ile Phe Gln Glu
                185                 190                 195
Leu Glu Asn Phe Ser Leu Glu Gln Glu Glu Glu Asp Val Pro Asp
                200                 205                 210
Gln Glu Gln Ser Ser Ser Ile Glu Thr Pro Ser Glu Glu Val Ala
                215                 220                 225
Ser Pro His Ser

<210> 5
<211> 360
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2367767CD1

<400> 5
Met Phe Val Ala Arg Ser Ile Ala Ala Asp His Lys Asp Leu Ile
 1              5                   10                  15
His Asp Val Ser Phe Asp Phe His Gly Arg Arg Met Ala Thr Cys
                20                  25                  30
Ser Ser Asp Gln Ser Val Lys Val Trp Asp Lys Ser Glu Ser Gly
                35                  40                  45
Asp Trp His Cys Thr Ala Ser Trp Lys Thr His Ser Gly Ser Val
                50                  55                  60
Trp Arg Val Thr Trp Ala His Pro Glu Phe Gly Gln Val Leu Ala
                65                  70                  75
Ser Cys Ser Phe Asp Arg Thr Ala Ala Val Trp Glu Glu Ile Val
                80                  85                  90
Gly Glu Ser Asn Asp Lys Leu Arg Gly Gln Ser His Trp Val Lys
                95                  100                 105
Arg Thr Thr Leu Val Asp Ser Arg Thr Ser Val Thr Asp Val Lys
                110                 115                 120
Phe Ala Pro Lys His Met Gly Leu Met Leu Ala Thr Cys Ser Ala
                125                 130                 135
Asp Gly Ile Val Arg Ile Tyr Glu Ala Pro Asp Val Met Asn Leu
                140                 145                 150
Ser Gln Trp Ser Leu Gln His Glu Ile Ser Cys Lys Leu Ser Cys
                155                 160                 165
Ser Cys Ile Ser Trp Asn Pro Ser Ser Ser Arg Ala His Ser Pro
                170                 175                 180
Met Ile Ala Val Gly Ser Asp Asp Ser Ser Pro Asn Ala Met Ala
                185                 190                 195
Lys Val Gln Ile Phe Glu Tyr Asn Glu Asn Thr Arg Lys Tyr Ala
                200                 205                 210
Lys Ala Glu Thr Leu Met Thr Val Thr Asp Pro Val His Asp Ile
                215                 220                 225
Ala Phe Ala Pro Asn Leu Gly Arg Ser Phe His Ile Leu Ala Ile
                230                 235                 240
Ala Thr Lys Asp Val Arg Ile Phe Thr Leu Lys Pro Val Arg Lys
```

```
                245                     250                     255
Glu Leu Thr Ser Ser Gly Gly Pro Thr Lys Phe Glu Ile His Ile
                260                     265                     270
Val Ala Gln Phe Asp Asn His Asn Ser Gln Val Trp Arg Val Ser
                275                     280                     285
Trp Asn Ile Thr Gly Thr Val Leu Ala Ser Ser Gly Asp Asp Gly
                290                     295                     300
Cys Val Arg Leu Trp Lys Ala Asn Tyr Met Asp Asn Trp Lys Cys
                305                     310                     315
Thr Gly Ile Leu Lys Gly Asn Gly Ser Pro Val Asn Gly Ser Ser
                320                     325                     330
Gln Gln Gly Thr Ser Asn Pro Ser Leu Gly Ser Asn Ile Pro Ser
                335                     340                     345
Leu Gln Asn Ser Leu Asn Gly Ser Ser Ala Gly Arg Lys His Ser
                350                     355                     360

<210> 6
<211> 460
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3090433CD1

<400> 6
Met Ala Asn Asp Pro Leu Glu Gly Phe His Glu Val Asn Leu Ala
 1               5                  10                  15
Ser Pro Thr Ser Pro Asp Leu Leu Gly Val Tyr Glu Ser Gly Thr
                20                  25                  30
Gln Glu Gln Thr Thr Ser Pro Ser Val Ile Tyr Arg Pro His Pro
                35                  40                  45
Ser Ala Leu Ser Ser Val Pro Ile Gln Ala Asn Ala Leu Asp Val
                50                  55                  60
Ser Glu Leu Pro Thr Gln Pro Val Tyr Ser Ser Pro Arg Arg Leu
                65                  70                  75
Asn Cys Ala Glu Ile Ser Ser Ile Ser Phe His Val Thr Asp Pro
                80                  85                  90
Ala Pro Cys Ser Thr Ser Gly Val Thr Ala Gly Leu Thr Lys Leu
                95                 100                 105
Thr Thr Arg Lys Asp Asn Tyr Asn Ala Glu Arg Glu Phe Leu Gln
               110                 115                 120
Gly Ala Thr Ile Thr Glu Ala Cys Asp Gly Ser Asp Asp Ile Phe
               125                 130                 135
Gly Leu Ser Thr Asp Ser Leu Ser Arg Leu Arg Ser Pro Ser Val
               140                 145                 150
Leu Glu Val Arg Glu Lys Gly Tyr Glu Arg Leu Lys Glu Glu Leu
               155                 160                 165
Ala Lys Ala Gln Arg Glu Leu Lys Leu Lys Asp Glu Glu Cys Glu
               170                 175                 180
Arg Leu Ser Lys Val Arg Asp Gln Leu Gly Gln Glu Leu Glu Glu
               185                 190                 195
Leu Thr Ala Ser Leu Phe Glu Glu Ala His Lys Met Val Arg Glu
               200                 205                 210
Ala Asn Ile Lys Gln Ala Thr Ala Glu Lys Gln Leu Lys Glu Ala
               215                 220                 225
Gln Gly Lys Ile Asp Val Leu Gln Ala Glu Val Ala Ala Leu Lys
               230                 235                 240
Thr Leu Val Leu Ser Ser Ser Pro Thr Ser Pro Thr Gln Glu Pro
               245                 250                 255
Leu Pro Gly Gly Lys Thr Pro Phe Lys Lys Gly His Thr Arg Asn
               260                 265                 270
Lys Ser Thr Ser Ser Ala Met Ser Gly Ser His Gln Asp Leu Ser
```

```
                275                     280                     285
Val Ile Gln Pro Ile Val Lys Asp Cys Lys Glu Ala Asp Leu Ser
                290                     295                     300
Leu Tyr Asn Glu Phe Arg Leu Trp Lys Asp Glu Pro Thr Met Asp
                305                     310                     315
Arg Thr Cys Pro Phe Leu Asp Lys Ile Tyr Gln Glu Asp Ile Phe
                320                     325                     330
Pro Cys Leu Thr Phe Ser Lys Ser Glu Leu Ala Ser Ala Val Leu
                335                     340                     345
Glu Ala Val Glu Asn Asn Thr Leu Ser Ile Glu Pro Val Gly Leu
                350                     355                     360
Gln Pro Ile Arg Phe Val Lys Ala Ser Ala Val Glu Cys Gly Gly
                365                     370                     375
Pro Lys Lys Cys Ala Leu Thr Gly Gln Ser Lys Ser Cys Lys His
                380                     385                     390
Arg Ile Lys Leu Gly Asp Ser Ser Asn Tyr Tyr Tyr Ile Ser Pro
                395                     400                     405
Phe Cys Arg Tyr Arg Ile Thr Ser Val Cys Asn Phe Phe Thr Tyr
                410                     415                     420
Ile Arg Tyr Ile Gln Gln Gly Leu Val Lys Gln Gln Asp Val Asp
                425                     430                     435
Gln Met Phe Trp Glu Val Met Gln Leu Arg Lys Glu Met Ser Leu
                440                     445                     450
Ala Lys Leu Gly Tyr Phe Lys Glu Glu Leu
                455                     460
```

<210> 7
<211> 239
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3800591CD1

<400> 7

```
Met Gln Asp Pro Asn Ala Asp Thr Glu Trp Asn Asp Ile Leu Arg
  1               5                  10                  15
Lys Lys Gly Ile Leu Pro Pro Lys Glu Ser Leu Lys Glu Leu Glu
                 20                  25                  30
Glu Glu Ala Glu Glu Glu Gln Arg Ile Leu Gln Gln Ser Val Val
                 35                  40                  45
Lys Thr Tyr Glu Asp Met Thr Leu Glu Glu Leu Glu Asp His Glu
                 50                  55                  60
Asp Glu Phe Asn Glu Glu Asp Glu Arg Ala Ile Glu Met Tyr Arg
                 65                  70                  75
Arg Arg Arg Leu Ala Glu Trp Lys Ala Thr Lys Leu Lys Asn Lys
                 80                  85                  90
Phe Gly Glu Val Leu Glu Ile Ser Gly Lys Asp Tyr Val Gln Glu
                 95                 100                 105
Val Thr Lys Ala Gly Glu Gly Leu Trp Val Ile Leu His Leu Tyr
                110                 115                 120
Lys Gln Gly Ile Pro Leu Cys Ala Leu Ile Asn Gln His Leu Ser
                125                 130                 135
Gly Leu Ala Arg Lys Phe Pro Asp Val Lys Phe Ile Lys Ala Ile
                140                 145                 150
Ser Thr Thr Cys Ile Pro Asn Tyr Pro Asp Arg Asn Leu Pro Thr
                155                 160                 165
Ile Phe Val Tyr Leu Glu Gly Asp Ile Lys Ala Gln Phe Ile Gly
                170                 175                 180
Pro Leu Val Phe Gly Gly Met Asn Leu Thr Arg Asp Glu Leu Glu
                185                 190                 195
Trp Lys Leu Ser Glu Ser Gly Ala Ile Met Thr Asp Leu Glu Glu
                200                 205                 210
```

Asn Pro Lys Lys Pro Ile Glu Asp Val Leu Leu Ser Ser Val Arg
215 220 225
Arg Ser Val Leu Met Lys Arg Asp Ser Asp Ser Glu Gly Asp
230 235

<210> 8
<211> 334
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5308471CD1

<400> 8
Met Arg Leu Thr Pro Arg Ala Leu Cys Ser Ala Ala Gln Ala Ala
1 5 10 15
Trp Arg Glu Asn Phe Pro Leu Cys Gly Arg Asp Val Ala Arg Trp
20 25 30
Phe Pro Gly His Met Ala Lys Gly Leu Lys Lys Met Gln Ser Ser
35 40 45
Leu Lys Leu Val Asp Cys Ile Ile Glu Val His Asp Ala Arg Ile
50 55 60
Pro Leu Ser Gly Arg Asn Pro Leu Phe Gln Glu Thr Leu Gly Leu
65 70 75
Lys Pro His Leu Leu Val Leu Asn Lys Met Asp Leu Ala Asp Leu
80 85 90
Thr Glu Gln Gln Lys Ile Met Gln His Leu Glu Gly Glu Gly Leu
95 100 105
Lys Asn Val Ile Phe Thr Asn Cys Val Lys Asp Glu Asn Val Lys
110 115 120
Gln Ile Ile Pro Met Val Thr Glu Leu Ile Gly Arg Ser His Arg
125 130 135
Tyr His Arg Lys Glu Asn Leu Glu Tyr Cys Ile Met Val Ile Gly
140 145 150
Val Pro Asn Val Gly Lys Ser Ser Leu Ile Asn Ser Leu Arg Arg
155 160 165
Gln His Leu Arg Lys Gly Lys Ala Thr Arg Val Gly Gly Glu Pro
170 175 180
Gly Ile Thr Arg Ala Val Met Ser Lys Ile Gln Val Ser Glu Arg
185 190 195
Pro Leu Met Phe Leu Leu Asp Thr Pro Gly Val Leu Ala Pro Arg
200 205 210
Ile Glu Ser Val Glu Thr Gly Leu Lys Leu Ala Leu Cys Gly Thr
215 220 225
Val Leu Asp His Leu Val Gly Glu Glu Thr Met Ala Asp Tyr Leu
230 235 240
Leu Tyr Thr Leu Asn Lys His Gln Arg Phe Gly Tyr Val Gln His
245 250 255
Tyr Gly Leu Gly Ser Ala Cys Asp Asn Val Glu Arg Val Leu Lys
260 265 270
Ser Val Ala Val Lys Leu Gly Lys Thr Gln Lys Val Lys Val Leu
275 280 285
Thr Gly Thr Gly Asn Val Asn Val Ile Gln Pro Asn Tyr Pro Ala
290 295 300
Ala Ala Arg Asp Phe Leu Gln Thr Phe Arg Arg Gly Leu Leu Gly
305 310 315
Ser Val Met Leu Asp Leu Asp Val Leu Arg Gly His Pro Pro Ala
320 325 330
Glu Thr Leu Pro

<210> 9
<211> 341
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5324322CD1

<400> 9

```
Met Glu Arg Ala Val Pro Leu Ala Val Pro Leu Gly Gln Thr Glu
 1               5                  10                  15
Val Phe Gln Ala Leu Gln Arg Leu His Met Thr Ile Phe Ser Gln
                20                  25                  30
Ser Val Ser Pro Cys Gly Lys Phe Leu Ala Ala Gly Asn Asn Tyr
                35                  40                  45
Gly Gln Ile Ala Ile Phe Ser Leu Ser Ser Ala Leu Ser Ser Glu
                50                  55                  60
Ala Lys Glu Glu Ser Lys Lys Pro Val Val Thr Phe Gln Ala His
                65                  70                  75
Asp Gly Pro Val Tyr Ser Met Val Ser Thr Asp Arg His Leu Leu
                80                  85                  90
Ser Ala Gly Asp Gly Glu Val Lys Ala Trp Leu Trp Ala Glu Met
                95                 100                 105
Leu Lys Lys Gly Cys Lys Glu Leu Trp Arg Arg Gln Pro Pro Tyr
               110                 115                 120
Arg Thr Ser Leu Glu Val Pro Glu Ile Asn Ala Leu Leu Leu Val
               125                 130                 135
Pro Lys Glu Asn Ser Leu Ile Leu Ala Gly Gly Asp Cys Gln Leu
               140                 145                 150
His Thr Met Asp Leu Glu Thr Gly Thr Phe Thr Arg Val Leu Arg
               155                 160                 165
Gly His Thr Asp Tyr Ile His Cys Leu Ala Leu Arg Glu Arg Ser
               170                 175                 180
Pro Glu Val Leu Ser Gly Gly Glu Asp Gly Ala Val Arg Leu Trp
               185                 190                 195
Asp Leu Arg Thr Ala Lys Glu Val Gln Thr Ile Glu Val Tyr Lys
               200                 205                 210
His Glu Glu Cys Ser Arg Pro His Asn Gly Arg Trp Ile Gly Cys
               215                 220                 225
Leu Ala Thr Asp Ser Asp Trp Met Val Cys Gly Gly Gly Pro Ala
               230                 235                 240
Leu Thr Leu Trp His Leu Arg Ser Ser Thr Pro Thr Thr Ile Phe
               245                 250                 255
Pro Ile Arg Ala Pro Gln Lys His Val Thr Phe Tyr Gln Asp Leu
               260                 265                 270
Ile Leu Ser Ala Gly Gln Gly Arg Cys Val Asn Gln Trp Gln Leu
               275                 280                 285
Ser Gly Glu Leu Lys Ala Gln Val Pro Gly Ser Ser Pro Gly Leu
               290                 295                 300
Leu Ser Leu Ser Leu Asn Gln Gln Pro Ala Ala Pro Glu Cys Lys
               305                 310                 315
Val Leu Thr Ala Ala Gly Asn Ser Cys Arg Val Asp Val Phe Thr
               320                 325                 330
Asn Leu Gly Tyr Arg Ala Phe Ser Leu Ser Phe
               335                 340
```

<210> 10
<211> 513
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 067184CD1

<400> 10

Met Ser Ile Glu Ile Glu Ser Ser Asp Val Ile Arg Leu Ile Met
1 5 10 15
Gln Tyr Leu Lys Glu Asn Ser Leu His Arg Ala Leu Ala Thr Leu
20 25 30
Gln Glu Glu Thr Thr Val Ser Leu Asn Thr Val Asp Ser Ile Glu
35 40 45
Ser Phe Val Ala Asp Ile Asn Ser Gly His Trp Asp Thr Val Leu
50 55 60
Gln Ala Ile Gln Ser Leu Lys Leu Pro Asp Lys Thr Leu Ile Asp
65 70 75
Leu Tyr Glu Gln Val Val Leu Glu Leu Ile Glu Leu Arg Glu Leu
80 85 90
Gly Ala Ala Arg Ser Leu Leu Arg Gln Thr Asp Pro Met Ile Met
95 100 105
Leu Lys Gln Thr Gln Pro Glu Arg Tyr Ile His Leu Glu Asn Leu
110 115 120
Leu Ala Arg Ser Tyr Phe Asp Pro Arg Glu Ala Tyr Pro Asp Gly
125 130 135
Ser Ser Lys Glu Lys Arg Arg Ala Ala Ile Ala Gln Ala Leu Ala
140 145 150
Gly Glu Val Ser Val Val Pro Pro Ser Arg Leu Met Ala Leu Leu
155 160 165
Gly Gln Ala Leu Lys Trp Gln Gln His Gln Gly Leu Leu Pro Pro
170 175 180
Gly Met Thr Ile Asp Leu Phe Arg Gly Lys Ala Ala Val Lys Asp
185 190 195
Val Glu Glu Glu Lys Phe Pro Thr Gln Leu Ser Arg His Ile Lys
200 205 210
Phe Gly Gln Lys Ser His Val Glu Cys Ala Arg Phe Ser Pro Asp
215 220 225
Gly Gln Tyr Leu Val Thr Gly Ser Val Asp Gly Phe Ile Glu Val
230 235 240
Trp Asn Phe Thr Thr Gly Lys Ile Arg Lys Asp Leu Lys Tyr Gln
245 250 255
Ala Gln Asp Asn Phe Met Met Met Asp Asp Ala Val Leu Cys Met
260 265 270
Cys Phe Ser Arg Asp Thr Glu Met Leu Ala Thr Gly Ala Gln Asp
275 280 285
Gly Lys Ile Lys Val Trp Lys Ile Gln Ser Gly Gln Cys Leu Arg
290 295 300
Arg Phe Glu Arg Ala His Ser Lys Gly Val Thr Cys Leu Ser Phe
305 310 315
Ser Lys Asp Ser Ser Gln Ile Leu Ser Ala Ser Phe Asp Gln Thr
320 325 330
Ile Arg Ile His Gly Leu Lys Ser Gly Lys Thr Leu Lys Glu Phe
335 340 345
Arg Gly His Ser Ser Phe Val Asn Glu Ala Thr Phe Thr Gln Asp
350 355 360
Gly His Tyr Ile Ile Ser Ala Ser Ser Asp Gly Thr Val Lys Ile
365 370 375
Trp Asn Met Lys Thr Thr Glu Cys Ser Asn Thr Phe Lys Ser Leu
380 385 390
Gly Ser Thr Ala Gly Thr Asp Ile Thr Val Asn Ser Val Ile Leu
395 400 405
Leu Pro Lys Asn Pro Glu His Phe Val Val Cys Asn Arg Ser Asn
410 415 420
Thr Val Val Ile Met Asn Met Gln Gly Gln Ile Val Arg Ser Phe
425 430 435
Ser Ser Gly Lys Arg Glu Gly Gly Asp Phe Val Cys Cys Ala Leu
440 445 450
Ser Pro Arg Gly Glu Trp Ile Tyr Cys Val Gly Glu Asp Phe Val
455 460 465
Leu Tyr Cys Phe Ser Thr Val Thr Gly Lys Leu Glu Arg Thr Leu

```
                470                 475                 480
Thr Val His Glu Lys Asp Val Ile Gly Ile Ala His His Pro His
                485                 490                 495
Gln Asn Leu Ile Ala Thr Tyr Ser Glu Asp Gly Leu Leu Lys Leu
                500                 505                 510
Trp Lys Pro
```

<210> 11
<211> 186
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 722896CD1

<400> 11

```
Met Ile Ala Leu Phe Asn Lys Leu Leu Asp Trp Phe Lys Ala Leu
  1             5                   10                  15
Phe Trp Lys Glu Glu Met Glu Leu Thr Leu Val Gly Leu Gln Tyr
                20                  25                  30
Ser Gly Lys Thr Thr Phe Val Asn Val Ile Ala Ser Gly Gln Phe
                35                  40                  45
Asn Glu Asp Met Ile Pro Thr Val Gly Phe Asn Met Arg Lys Ile
                50                  55                  60
Thr Lys Gly Asn Val Thr Ile Lys Leu Trp Asp Ile Gly Gly Gln
                65                  70                  75
Pro Arg Phe Arg Ser Met Trp Glu Arg Tyr Cys Arg Gly Val Ser
                80                  85                  90
Ala Ile Val Tyr Met Val Asp Ala Ala Asp Gln Glu Lys Ile Glu
                95                  100                 105
Ala Ser Lys Asn Glu Leu His Asn Leu Leu Asp Lys Pro Gln Leu
                110                 115                 120
Gln Gly Ile Pro Val Leu Val Leu Gly Asn Lys Arg Asp Leu Pro
                125                 130                 135
Gly Ala Leu Asp Glu Lys Glu Leu Ile Glu Lys Met Asn Leu Ser
                140                 145                 150
Ala Ile Gln Asp Arg Glu Ile Cys Cys Tyr Ser Ile Ser Cys Lys
                155                 160                 165
Glu Lys Asp Asn Ile Asp Ile Thr Leu Gln Trp Leu Ile Gln His
                170                 175                 180
Ser Lys Ser Arg Arg Ser
                185
```

<210> 12
<211> 204
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1571739CD1

<400> 12

```
Met Asn Asp Val Lys Leu Ala Val Leu Gly Gly Glu Gly Thr Gly
  1             5                   10                  15
Lys Ser Ala Leu Thr Val Arg Phe Leu Thr Lys Arg Phe Ile Gly
                20                  25                  30
Glu Tyr Ala Ser Asn Phe Glu Ser Ile Tyr Lys Lys His Leu Cys
                35                  40                  45
Leu Glu Arg Lys Gln Leu Asn Leu Glu Ile Tyr Asp Pro Cys Ser
                50                  55                  60
Gln Thr Gln Lys Ala Lys Phe Ser Leu Thr Ser Glu Leu His Trp
                65                  70                  75
```

```
Ala Asp Gly Phe Val Ile Val Tyr Asp Ile Ser Asp Arg Ser Ser
                80                  85                  90
Phe Ala Phe Ala Lys Ala Leu Ile Tyr Arg Ile Arg Glu Pro Gln
                95                  100                 105
Thr Ser His Cys Lys Arg Ala Val Glu Ser Ala Val Phe Leu Val
                110                 115                 120
Gly Asn Lys Arg Asp Leu Cys His Val Arg Glu Val Gly Trp Glu
                125                 130                 135
Glu Gly Gln Lys Leu Ala Leu Glu Asn Arg Cys Gln Phe Cys Glu
                140                 145                 150
Leu Ser Ala Ala Glu Gln Ser Leu Glu Val Glu Met Met Phe Ile
                155                 160                 165
Arg Ile Ile Lys Asp Ile Leu Ile Asn Phe Lys Leu Lys Glu Lys
                170                 175                 180
Arg Arg Pro Ser Gly Ser Lys Ser Met Ala Lys Leu Ile Asn Asn
                185                 190                 195
Val Phe Gly Lys Arg Arg Lys Ser Val
                200
<210> 13
<211> 100
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1739479CD1

<400> 13
Met Trp Asp Ser Lys Lys Ile Gly Leu Arg Gln His His Cys Arg
 1               5                  10                  15
Lys Cys Gly Lys Ala Val Cys Gly Lys Cys Ser Ser Lys Arg Ser
                20                  25                  30
Ser Ile Pro Leu Met Gly Phe Glu Phe Glu Val Arg Val Cys Asp
                35                  40                  45
Ser Cys His Glu Ala Ile Thr Asp Glu Glu Arg Ala Pro Thr Ala
                50                  55                  60
Thr Phe His Asp Ser Lys His Asn Ile Val His Val His Phe Asp
                65                  70                  75
Ala Thr Arg Gly Trp Leu Leu Thr Ser Gly Thr Asp Lys Val Ile
                80                  85                  90
Lys Leu Trp Asp Met Thr Pro Val Val Ser
                95                  100
<210> 14
<211> 795
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1999147CD1

<400> 14
Met Thr Ser Gly Ala Thr Arg Tyr Arg Leu Ser Cys Ser Leu Arg
 1               5                  10                  15
Gly His Glu Leu Asp Val Arg Gly Leu Val Cys Cys Ala Tyr Pro
                20                  25                  30
Pro Gly Ala Phe Val Ser Val Ser Arg Asp Arg Thr Thr Arg Leu
                35                  40                  45
Trp Ala Pro Asp Ser Pro Asn Arg Ser Phe Thr Glu Met His Cys
                50                  55                  60
Met Ser Gly His Ser Asn Phe Val Ser Cys Val Cys Ile Ile Pro
                65                  70                  75
Ser Ser Asp Ile Tyr Pro His Gly Leu Ile Ala Thr Gly Gly Asn
```

```
                    80                  85                  90
Asp His Asn Ile Cys Ile Phe Ser Leu Asp Ser Pro Met Pro Leu
                    95                  100                 105
Tyr Ile Leu Lys Gly His Lys Asn Thr Val Cys Ser Leu Ser Ser
                    110                 115                 120
Gly Lys Phe Gly Thr Leu Leu Ser Gly Ser Trp Asp Thr Thr Ala
                    125                 130                 135
Lys Val Trp Leu Asn Asp Lys Cys Met Met Thr Leu Gln Gly His
                    140                 145                 150
Thr Ala Ala Val Trp Ala Val Lys Ile Leu Pro Glu Gln Gly Leu
                    155                 160                 165
Met Leu Thr Gly Ser Ala Asp Lys Thr Val Lys Leu Trp Lys Ala
                    170                 175                 180
Gly Arg Cys Glu Arg Thr Phe Ser Gly His Glu Asp Cys Val Arg
                    185                 190                 195
Gly Leu Ala Ile Leu Ser Glu Thr Glu Phe Leu Ser Cys Ala Asn
                    200                 205                 210
Asp Ala Ser Ile Arg Arg Trp Gln Ile Thr Gly Glu Cys Leu Glu
                    215                 220                 225
Val Tyr Tyr Gly His Thr Asn Tyr Ile Tyr Ser Ile Ser Val Phe
                    230                 235                 240
Pro Asn Cys Arg Asp Phe Val Thr Thr Ala Glu Asp Arg Ser Leu
                    245                 250                 255
Arg Ile Trp Lys His Gly Glu Cys Ala Gln Thr Ile Arg Leu Pro
                    260                 265                 270
Ala Gln Ser Ile Trp Cys Cys Cys Val Leu Asp Asn Gly Asp Ile
                    275                 280                 285
Val Val Gly Ala Ser Asp Gly Ile Ile Arg Val Phe Thr Glu Ser
                    290                 295                 300
Glu Asp Arg Thr Ala Ser Ala Glu Glu Ile Lys Ala Phe Glu Lys
                    305                 310                 315
Glu Leu Ser His Ala Thr Ile Asp Ser Lys Thr Gly Asp Leu Gly
                    320                 325                 330
Asp Ile Asn Ala Glu Gln Leu Pro Gly Arg Glu His Leu Asn Glu
                    335                 340                 345
Pro Gly Thr Arg Glu Gly Gln Thr Arg Leu Ile Arg Asp Gly Glu
                    350                 355                 360
Lys Val Glu Ala Tyr Gln Trp Ser Val Ser Glu Gly Arg Trp Ile
                    365                 370                 375
Lys Ile Gly Asp Val Val Gly Ser Ser Gly Ala Asn Gln Gln Thr
                    380                 385                 390
Ser Gly Lys Val Leu Tyr Glu Gly Lys Glu Phe Asp Tyr Val Phe
                    395                 400                 405
Ser Ile Asp Val Asn Glu Gly Gly Pro Ser Tyr Lys Leu Pro Tyr
                    410                 415                 420
Asn Thr Ser Asp Asp Pro Trp Leu Thr Ala Tyr Asn Phe Leu Gln
                    425                 430                 435
Lys Asn Asp Leu Asn Pro Met Phe Leu Asp Gln Val Ala Lys Phe
                    440                 445                 450
Ile Ile Asp Asn Thr Lys Gly Gln Met Leu Gly Leu Gly Asn Pro
                    455                 460                 465
Ser Phe Ser Asp Pro Phe Thr Gly Gly Gly Arg Tyr Val Pro Gly
                    470                 475                 480
Ser Ser Gly Ser Ser Asn Thr Leu Pro Thr Ala Asp Pro Phe Thr
                    485                 490                 495
Gly Ala Gly Arg Tyr Val Pro Gly Ser Ala Ser Met Gly Thr Thr
                    500                 505                 510
Met Ala Gly Val Asp Pro Phe Thr Gly Asn Ser Ala Tyr Arg Ser
                    515                 520                 525
Ala Ala Ser Lys Thr Met Asn Ile Tyr Phe Pro Lys Lys Glu Ala
                    530                 535                 540
Val Thr Phe Asp Gln Ala Asn Pro Thr Gln Ile Leu Gly Lys Leu
                    545                 550                 555
```

Lys Glu Leu Asn Gly Thr Ala Pro Glu Glu Lys Lys Leu Thr Glu
560 565 570
Asp Asp Leu Ile Leu Leu Glu Lys Ile Leu Ser Leu Ile Cys Asn
575 580 585
Ser Ser Ser Glu Lys Pro Thr Val Gln Gln Leu Gln Ile Leu Trp
590 595 600
Lys Ala Ile Asn Cys Pro Glu Asp Ile Val Phe Pro Ala Leu Asp
605 610 615
Ile Leu Arg Leu Ser Ile Lys His Pro Ser Val Asn Glu Asn Phe
620 625 630
Cys Asn Glu Lys Glu Gly Ala Gln Phe Ser Ser His Leu Ile Asn
635 640 645
Leu Leu Asn Pro Lys Gly Lys Pro Ala Asn Gln Leu Leu Ala Leu
650 655 660
Arg Thr Phe Cys Asn Cys Phe Val Gly Gln Ala Gly Gln Lys Leu
665 670 675
Met Met Ser Gln Arg Glu Ser Leu Met Ser His Ala Ile Glu Leu
680 685 690
Lys Ser Gly Ser Asn Lys Asn Ile His Ile Ala Leu Ala Thr Leu
695 700 705
Ala Leu Asn Tyr Ser Val Cys Phe His Lys Asp His Asn Ile Glu
710 715 720
Gly Lys Ala Gln Cys Leu Ser Leu Ile Ser Thr Ile Leu Glu Val
725 730 735
Val Gln Asp Leu Glu Ala Thr Phe Arg Leu Leu Val Ala Leu Gly
740 745 750
Thr Leu Ile Ser Asp Asp Ser Asn Ala Val Gln Leu Ala Lys Ser
755 760 765
Leu Gly Val Asp Ser Gln Ile Lys Lys Tyr Ser Ser Val Ser Glu
770 775 780
Pro Ala Lys Val Ser Glu Cys Cys Arg Phe Ile Leu Asn Leu Leu
785 790 795

<210> 15
<211> 393
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2182085CD1

<400> 15
Met Glu Asp Phe Glu Asp Asp Pro Arg Ala Leu Gly Ala Arg Gly
1 5 10 15
His Arg Arg Ser Val Ser Arg Gly Ser Tyr Gln Leu Gln Ala Gln
20 25 30
Met Asn Arg Ala Val Tyr Glu Asp Arg Pro Pro Gly Ser Val Val
35 40 45
Pro Thr Ser Ala Ala Glu Ala Ser Arg Ala Met Ala Gly Asp Thr
50 55 60
Ser Leu Ser Glu Asn Tyr Ala Phe Ala Gly Met Tyr His Val Phe
65 70 75
Asp Gln His Val Asp Glu Ala Val Pro Arg Val Arg Phe Ala Asn
80 85 90
Asp Asp Arg His Arg Leu Ala Cys Cys Ser Leu Asp Gly Ser Ile
95 100 105
Ser Leu Cys Gln Leu Val Pro Ala Pro Pro Thr Val Leu Arg Val
110 115 120
Leu Arg Gly His Thr Arg Gly Val Ser Asp Phe Ala Trp Ser Leu
125 130 135
Ser Asn Asp Ile Leu Val Ser Thr Ser Leu Asp Ala Thr Met Arg
140 145 150

Ile Trp Ala Ser Glu Asp Gly Arg Cys Ile Arg Glu Ile Pro Asp
155 160 165
Pro Asp Ser Ala Glu Leu Leu Cys Cys Thr Phe Gln Pro Val Asn
170 175 180
Asn Asn Leu Thr Val Val Gly Asn Ala Lys His Asn Val His Val
185 190 195
Met Asn Ile Ser Thr Gly Lys Lys Val Lys Gly Gly Ser Ser Lys
200 205 210
Leu Thr Gly Arg Val Leu Ala Leu Ser Phe Asp Ala Pro Gly Arg
215 220 225
Leu Leu Trp Ala Gly Asp Asp Arg Gly Ser Val Phe Ser Phe Leu
230 235 240
Phe Asp Met Ala Thr Gly Lys Leu Thr Lys Ala Lys Arg Leu Val
245 250 255
Val His Glu Gly Ser Pro Val Thr Ser Ile Ser Ala Arg Ser Trp
260 265 270
Val Ser Arg Glu Ala Arg Asp Pro Ser Leu Leu Ile Asn Ala Cys
275 280 285
Leu Asn Lys Leu Leu Leu Tyr Arg Val Val Asp Asn Glu Gly Thr
290 295 300
Leu Gln Leu Lys Arg Ser Phe Pro Ile Glu Gln Ser Ser His Pro
305 310 315
Val Arg Ser Ile Phe Cys Pro Leu Met Ser Phe Arg Gln Gly Ala
320 325 330
Cys Val Val Thr Gly Ser Glu Asp Met Cys Val His Phe Phe Asp
335 340 345
Val Glu Arg Ala Ala Lys Ala Ala Val Asn Lys Leu Gln Gly His
350 355 360
Ser Ala Pro Val Leu Asp Val Ser Phe Asn Cys Asp Glu Ser Leu
365 370 375
Leu Ala Ser Ser Asp Ala Ser Gly Met Val Ile Val Trp Arg Arg
380 385 390
Glu Gln Lys

<210> 16
<211> 485
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2216640CD1

<400> 16
Met Ala Ala Ala Val Ala Asp Glu Ala Val Ala Arg Asp Val Gln
1 5 10 15
Arg Leu Leu Val Gln Phe Gln Asp Glu Gly Gly Gln Leu Leu Gly
20 25 30
Ser Pro Phe Asp Val Pro Val Asp Ile Thr Pro Asp Arg Leu Gln
35 40 45
Leu Val Cys Asn Ala Leu Leu Ala Gln Glu Asp Pro Leu Pro Leu
50 55 60
Ala Phe Phe Val His Asp Ala Glu Ile Val Ser Ser Leu Gly Lys
65 70 75
Thr Leu Glu Ser Gln Ala Val Glu Thr Glu Lys Val Leu Asp Ile
80 85 90
Ile Tyr Gln Pro Gln Ala Ile Phe Arg Val Arg Ala Val Thr Arg
95 100 105
Cys Thr Ser Ser Leu Glu Gly His Ser Glu Ala Val Ile Ser Val
110 115 120
Ala Phe Ser Pro Thr Gly Lys Tyr Leu Ala Ser Gly Ser Gly Asp
125 130 135
Thr Thr Val Arg Phe Trp Asp Leu Ser Thr Glu Thr Pro His Phe

```
                140                      145                      150
Thr Cys Lys Gly His Arg His Trp Val Leu Ser Ile Ser Trp Ser
                155                      160                      165
Pro Asp Gly Lys Lys Leu Ala Ser Gly Cys Lys Asn Gly Gln Ile
                170                      175                      180
Leu Leu Trp Asp Pro Ser Thr Gly Lys Gln Val Gly Arg Thr Leu
                185                      190                      195
Ala Gly His Ser Lys Trp Ile Thr Gly Leu Ser Trp Glu Pro Leu
                200                      205                      210
His Ala Asn Pro Glu Cys Arg Tyr Val Ala Ser Ser Ser Lys Asp
                215                      220                      225
Gly Ser Val Arg Ile Trp Asp Thr Thr Ala Gly Arg Cys Glu Arg
                230                      235                      240
Ile Leu Thr Gly His Thr Gln Ser Val Thr Cys Leu Arg Trp Gly
                245                      250                      255
Gly Asp Gly Leu Leu Tyr Ser Ala Ser Gln Asp Arg Thr Ile Lys
                260                      265                      270
Val Trp Arg Ala His Asp Gly Val Leu Cys Arg Thr Leu Gln Gly
                275                      280                      285
His Gly His Trp Val Asn Thr Met Ala Leu Ser Thr Asp Tyr Ala
                290                      295                      300
Leu Arg Thr Gly Ala Phe Glu Pro Ala Glu Ala Ser Val Asn Pro
                305                      310                      315
Gln Asp Leu Gln Gly Ser Leu Gln Glu Leu Lys Glu Arg Ala Leu
                320                      325                      330
Ser Arg Tyr Asn Leu Val Arg Gly Gln Gly Pro Glu Arg Leu Val
                335                      340                      345
Ser Gly Ser Asp Asp Phe Thr Leu Phe Leu Trp Ser Pro Ala Glu
                350                      355                      360
Asp Lys Lys Pro Leu Thr Arg Met Thr Gly His Gln Ala Leu Ile
                365                      370                      375
Asn Gln Val Leu Phe Ser Pro Asp Ser Arg Ile Val Ala Ser Ala
                380                      385                      390
Ser Phe Asp Lys Ser Ile Lys Leu Trp Asp Gly Arg Thr Gly Lys
                395                      400                      405
Tyr Leu Ala Ser Leu Arg Gly His Val Ala Ala Val Tyr Gln Ile
                410                      415                      420
Ala Trp Ser Ala Asp Ser Arg Leu Leu Val Ser Gly Ser Ser Asp
                425                      430                      435
Ser Thr Leu Lys Val Trp Asp Val Lys Ala Gln Lys Leu Ala Met
                440                      445                      450
Asp Leu Pro Gly His Ala Asp Glu Val Tyr Ala Val Asp Trp Ser
                455                      460                      465
Pro Asp Gly Gln Arg Val Ala Ser Gly Gly Lys Asp Lys Cys Leu
                470                      475                      480
Arg Ile Trp Arg Arg
                485
<210> 17
<211> 199
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2417361CD1

<400> 17
Met Asn Pro Arg Lys Lys Val Asp Leu Lys Leu Ile Ile Val Gly
  1               5                       10                       15
Ala Ile Gly Val Gly Lys Thr Ser Leu Leu His Gln Tyr Val His
                 20                       25                       30
Lys Thr Phe Tyr Glu Glu Tyr Gln Thr Thr Leu Gly Ala Ser Ile
                 35                       40                       45
```

Leu Ser Lys Ile Ile Ile Leu Gly Asp Thr Thr Leu Lys Leu Gln
50 55 60
Ile Trp Asp Thr Gly Gly Gln Glu Arg Phe Arg Ser Met Val Ser
65 70 75
Thr Phe Tyr Lys Gly Ser Asp Gly Cys Ile Leu Ala Phe Asp Val
80 85 90
Thr Asp Leu Glu Ser Phe Glu Ala Leu Asp Ile Trp Arg Gly Asp
95 100 105
Val Leu Ala Lys Ile Val Pro Met Glu Gln Ser Tyr Pro Met Val
110 115 120
Leu Leu Gly Asn Lys Ile Asp Leu Ala Asp Arg Lys Val Pro Gln
125 130 135
Glu Val Ala Gln Gly Trp Cys Arg Glu Lys Asp Ile Pro Tyr Phe
140 145 150
Glu Val Ser Ala Lys Asn Asp Ile Asn Val Val Gln Ala Phe Glu
155 160 165
Met Leu Ala Ser Arg Ala Leu Ser Arg Tyr Gln Ser Ile Leu Glu
170 175 180
Asn His Leu Thr Glu Ser Ile Lys Leu Ser Pro Asp Gln Ser Arg
185 190 195
Ser Arg Cys Cys

<210> 18
<211> 163
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2454384CD1

<400> 18
Met Glu Gly Pro Ser Leu Arg Gly Pro Ala Leu Arg Leu Ala Gly
1 5 10 15
Leu Pro Thr Gln Gln Asp Cys Asn Ile Gln Glu Lys Ile Asp Leu
20 25 30
Glu Ile Arg Met Arg Glu Gly Ile Trp Lys Leu Leu Ser Leu Ser
35 40 45
Thr Gln Lys Asp Gln Val Leu His Ala Val Lys Asn Leu Met Val
50 55 60
Cys Asn Ala Arg Leu Met Ala Tyr Thr Ser Glu Leu Gln Lys Leu
65 70 75
Glu Glu Gln Ile Ala Asn Gln Thr Gly Arg Cys Asp Val Lys Phe
80 85 90
Glu Ser Lys Glu Arg Thr Ala Cys Lys Gly Lys Ile Ala Ile Ser
95 100 105
Asp Ile Arg Ile Pro Leu Met Trp Lys Asp Ser Asp His Phe Ser
110 115 120
Asn Lys Glu Arg Ser Arg Arg Tyr Ala Ile Phe Cys Leu Phe Lys
125 130 135
Met Gly Ala Asn Val Phe Asp Thr Asp Val Val Asn Val Asp Lys
140 145 150
Thr Ile Thr Asp Ile Cys Phe Glu Asn Val Thr Ile Leu
155 160

<210> 19
<211> 290
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2610262CD1

<400> 19
Met Ala Ala Glu Ile Gln Pro Lys Pro Leu Thr Arg Lys Pro Ile
1 5 10 15
Leu Leu Gln Arg Met Glu Gly Ser Gln Glu Val Val Asn Met Ala
20 25 30
Val Ile Val Pro Lys Glu Glu Gly Val Ile Ser Val Ser Glu Asp
35 40 45
Arg Thr Val Arg Val Trp Leu Lys Arg Asp Ser Gly Gln Tyr Trp
50 55 60
Pro Ser Val Tyr His Ala Met Pro Ser Pro Cys Ser Cys Met Ser
65 70 75
Phe Asn Pro Glu Thr Arg Arg Leu Ser Ile Gly Leu Asp Asn Gly
80 85 90
Thr Ile Ser Glu Phe Ile Leu Ser Glu Asp Tyr Asn Lys Met Thr
95 100 105
Pro Val Lys Asn Tyr Gln Ala His Gln Ser Arg Val Thr Met Ile
110 115 120
Leu Phe Val Leu Glu Leu Glu Trp Val Leu Ser Thr Gly Gln Asp
125 130 135
Lys Gln Phe Ala Trp His Cys Ser Glu Ser Gly Gln Arg Leu Gly
140 145 150
Gly Tyr Arg Thr Ser Ala Val Ala Ser Gly Leu Gln Phe Asp Val
155 160 165
Glu Thr Arg His Val Phe Ile Gly Asp His Ser Gly Gln Val Thr
170 175 180
Ile Leu Lys Leu Glu Gln Glu Asn Cys Thr Leu Val Thr Thr Phe
185 190 195
Arg Gly His Thr Gly Gly Val Thr Ala Leu Cys Trp Asp Pro Val
200 205 210
Gln Arg Val Leu Phe Ser Gly Ser Ser Asp His Ser Val Ile Met
215 220 225
Trp Asp Ile Gly Gly Arg Lys Gly Thr Ala Ile Glu Leu Gln Gly
230 235 240
His Asn Asp Arg Val Gln Ala Leu Ser Tyr Ala Gln His Thr Arg
245 250 255
Gln Leu Ile Ser Cys Gly Gly Asp Gly Gly Ile Val Val Trp Asn
260 265 270
Met Asp Val Glu Arg Gln Glu Pro Leu Trp Ser Cys Phe Val Val
275 280 285
Met Ile Ser Ala Val
290

<210> 20
<211> 705
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2700075CD1

<400> 20
Met Gly Thr Trp Glu His Leu Val Ser Thr Gly Phe Asn Gln Met
1 5 10 15
Arg Glu Arg Glu Val Lys Leu Trp Asp Thr Arg Phe Phe Ser Ser
20 25 30
Ala Leu Ala Ser Leu Thr Leu Asp Thr Ser Leu Gly Cys Leu Val
35 40 45
Pro Leu Leu Asp Pro Asp Ser Gly Leu Leu Val Leu Ala Gly Lys
50 55 60
Gly Glu Arg Gln Leu Tyr Cys Tyr Glu Val Val Pro Gln Gln Pro
65 70 75
Ala Leu Ser Pro Val Thr Gln Cys Val Leu Glu Ser Val Leu Arg
80 85 90

Gly Ala Ala Leu Val Pro Arg Gln Ala Leu Ala Val Met Ser Cys
95 100 105
Glu Val Leu Arg Val Leu Gln Leu Ser Asp Thr Ala Ile Val Pro
110 115 120
Ile Gly Tyr His Val Pro Arg Lys Ala Val Glu Phe His Glu Asp
125 130 135
Leu Phe Pro Asp Thr Ala Gly Cys Val Pro Ala Thr Asp Pro His
140 145 150
Ser Trp Trp Ala Gly Asp Asn Gln Gln Val Gln Lys Val Ser Leu
155 160 165
Asn Pro Ala Cys Arg Pro His Pro Ser Phe Thr Ser Cys Leu Val
170 175 180
Pro Pro Ala Glu Pro Leu Pro Asp Thr Ala Gln Pro Ala Val Met
185 190 195
Glu Thr Pro Val Gly Asp Ala Asp Ala Ser Glu Gly Phe Ser Ser
200 205 210
Pro Pro Ser Ser Leu Thr Ser Pro Ser Thr Pro Ser Ser Leu Gly
215 220 225
Pro Ser Leu Ser Ser Thr Ser Gly Ile Gly Thr Ser Pro Ser Leu
230 235 240
Arg Ser Leu Gln Ser Leu Leu Gly Pro Ser Ser Lys Phe Arg His
245 250 255
Ala Gln Gly Thr Val Leu His Arg Asp Ser His Ile Thr Asn Leu
260 265 270
Lys Gly Leu Asn Leu Thr Thr Pro Gly Glu Ser Asp Gly Phe Cys
275 280 285
Ala Asn Lys Leu Arg Val Ala Val Pro Leu Leu Ser Ser Gly Gly
290 295 300
Gln Val Ala Val Leu Glu Leu Arg Lys Pro Gly Arg Leu Pro Asp
305 310 315
Thr Ala Leu Pro Thr Leu Gln Asn Gly Ala Ala Val Thr Asp Leu
320 325 330
Ala Trp Asp Pro Phe Asp Pro His Arg Leu Ala Val Ala Gly Glu
335 340 345
Asp Ala Arg Ile Arg Leu Trp Arg Val Pro Ala Glu Gly Leu Glu
350 355 360
Glu Val Leu Thr Thr Pro Glu Thr Val Leu Thr Gly His Thr Glu
365 370 375
Lys Ile Cys Ser Leu Arg Phe His Pro Leu Ala Ala Asn Val Leu
380 385 390
Ala Ser Ser Ser Tyr Asp Leu Thr Val Arg Ile Trp Asp Leu Gln
395 400 405
Ala Gly Ala Asp Arg Leu Lys Leu Gln Gly His Gln Asp Gln Ile
410 415 420
Phe Ser Leu Ala Trp Ser Pro Asp Gly Gln Gln Leu Ala Thr Val
425 430 435
Cys Lys Asp Gly Arg Val Arg Val Tyr Arg Pro Arg Ser Gly Pro
440 445 450
Glu Pro Leu Gln Glu Gly Pro Gly Pro Lys Gly Gly Arg Gly Ala
455 460 465
Arg Ile Val Trp Val Cys Asp Gly Arg Cys Leu Leu Val Ser Gly
470 475 480
Phe Asp Ser Gln Ser Glu Arg Gln Leu Leu Leu Tyr Glu Ala Glu
485 490 495
Ala Leu Ala Gly Gly Pro Leu Ala Val Leu Gly Leu Asp Val Ala
500 505 510
Pro Ser Thr Leu Leu Pro Ser Tyr Asp Pro Asp Thr Gly Leu Val
515 520 525
Leu Leu Thr Gly Lys Gly Asp Thr Arg Val Phe Leu Tyr Glu Leu
530 535 540
Leu Pro Glu Ser Pro Phe Phe Leu Glu Cys Asn Ser Phe Thr Ser
545 550 555
Pro Asp Pro His Lys Gly Leu Val Leu Leu Pro Lys Thr Glu Cys

```
                560                     565                     570
Asp Val Arg Glu Val Glu Leu Met Arg Cys Leu Arg Leu Arg Gln
                575                     580                     585
Ser Ser Leu Glu Pro Val Ala Phe Arg Leu Pro Arg Val Arg Lys
                590                     595                     600
Glu Phe Phe Gln Asp Asp Val Phe Pro Asp Thr Ala Val Ile Trp
                605                     610                     615
Glu Pro Val Leu Ser Ala Glu Ala Trp Leu Gln Gly Ala Asn Gly
                620                     625                     630
Gln Pro Trp Leu Leu Ser Leu Gln Pro Pro Asp Met Ser Pro Val
                635                     640                     645
Ser Gln Ala Pro Arg Glu Ala Pro Ala Arg Arg Ala Pro Ser Ser
                650                     655                     660
Ala Gln Tyr Leu Glu Glu Lys Ser Asp Gln Gln Lys Lys Glu Glu
                665                     670                     675
Leu Leu Asn Ala Met Val Ala Lys Leu Gly Asn Arg Glu Asp Pro
                680                     685                     690
Leu Pro Gln Asp Ser Phe Glu Gly Val Asp Glu Asp Glu Trp Asp
                695                     700                     705

<210> 21
<211> 454
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2786701CD1

<400> 21
Met Ala Ser Ser Glu Val Ala Arg His Leu Leu Phe Gln Ser His
 1               5                  10                  15
Met Ala Thr Lys Thr Thr Cys Met Ser Ser Gln Gly Ser Asp Asp
                20                  25                  30
Glu Gln Ile Lys Arg Glu Asn Ile Arg Ser Leu Thr Met Ser Gly
                35                  40                  45
His Val Gly Phe Glu Ser Leu Pro Asp Gln Leu Val Asn Arg Ser
                50                  55                  60
Ile Gln Gln Gly Phe Cys Phe Asn Ile Leu Cys Val Gly Glu Thr
                65                  70                  75
Gly Ile Gly Lys Ser Thr Leu Ile Asp Thr Leu Phe Asn Thr Asn
                80                  85                  90
Phe Glu Asp Tyr Glu Ser Ser His Phe Cys Pro Asn Val Lys Leu
                95                 100                 105
Lys Ala Gln Thr Tyr Glu Leu Gln Glu Ser Asn Val Gln Leu Lys
               110                 115                 120
Leu Thr Ile Val Asn Thr Val Gly Phe Gly Asp Gln Ile Asn Lys
               125                 130                 135
Glu Glu Ser Tyr Gln Pro Ile Val Asp Tyr Ile Asp Ala Gln Phe
               140                 145                 150
Glu Ala Tyr Leu Gln Glu Glu Leu Lys Ile Lys Arg Ser Leu Phe
               155                 160                 165
Thr Tyr His Asp Ser Arg Ile His Val Cys Leu Tyr Phe Ile Ser
               170                 175                 180
Pro Thr Gly His Ser Leu Lys Thr Leu Asp Leu Leu Thr Met Lys
               185                 190                 195
Asn Leu Asp Ser Lys Val Asn Ile Ile Pro Val Ile Ala Lys Ala
               200                 205                 210
Asp Thr Val Ser Lys Thr Glu Leu Gln Lys Phe Lys Ile Lys Leu
               215                 220                 225
Met Ser Glu Leu Val Ser Asn Gly Val Gln Ile Tyr Gln Phe Pro
               230                 235                 240
Thr Asp Asp Asp Thr Ile Ala Lys Val Asn Ala Ala Met Asn Gly
```

```
                    245                     250                     255
Gln Leu Pro Phe Ala Val Val Gly Ser Met Asp Glu Val Lys Val
                    260                     265                     270
Gly Asn Lys Met Val Lys Ala Arg Gln Tyr Pro Trp Gly Val Val
                    275                     280                     285
Gln Val Glu Asn Glu Asn His Cys Asp Phe Val Lys Leu Arg Glu
                    290                     295                     300
Met Leu Ile Cys Thr Asn Met Glu Asp Leu Arg Glu Gln Thr His
                    305                     310                     315
Thr Arg His Tyr Glu Leu Tyr Arg Arg Cys Lys Leu Glu Glu Met
                    320                     325                     330
Gly Phe Thr Asp Val Gly Pro Glu Asn Lys Pro Val Ser Val Gln
                    335                     340                     345
Glu Thr Tyr Glu Ala Lys Arg His Glu Phe His Gly Glu Arg Gln
                    350                     355                     360
Arg Lys Glu Glu Glu Met Lys Gln Met Phe Val Gln Arg Val Lys
                    365                     370                     375
Glu Lys Glu Ala Ile Leu Lys Glu Ala Glu Arg Glu Leu Gln Ala
                    380                     385                     390
Lys Phe Glu His Leu Lys Arg Leu His Gln Glu Glu Arg Met Lys
                    395                     400                     405
Leu Glu Glu Lys Arg Arg Leu Leu Glu Glu Glu Ile Ile Ala Phe
                    410                     415                     420
Ser Lys Lys Lys Ala Thr Ser Glu Ile Phe His Ser Gln Ser Phe
                    425                     430                     435
Leu Ala Thr Gly Ser Asn Leu Arg Lys Asp Lys Asp Arg Lys Asn
                    440                     445                     450
Ser Asn Phe Leu

<210> 22
<211> 433
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3068538CD1

<400> 22
Met Ala Gly Gln Asp Pro Ala Leu Ser Thr Ser His Pro Phe Tyr
 1               5                  10                  15
Asp Val Ala Arg His Gly Ile Leu Gln Val Ala Gly Asp Asp Arg
                20                  25                  30
Phe Gly Arg Arg Val Val Thr Phe Ser Cys Cys Arg Met Pro Pro
                35                  40                  45
Ser His Glu Leu Asp His Gln Arg Leu Leu Glu Tyr Leu Lys Tyr
                50                  55                  60
Thr Leu Asp Gln Tyr Val Glu Asn Asp Tyr Thr Ile Val Tyr Phe
                65                  70                  75
His Tyr Gly Leu Asn Ser Arg Asn Lys Pro Ser Leu Gly Trp Leu
                80                  85                  90
Gln Ser Ala Tyr Lys Glu Phe Asp Arg Lys Tyr Lys Lys Asn Leu
                95                 100                 105
Lys Ala Leu Tyr Val Val His Pro Thr Ser Phe Ile Lys Val Leu
               110                 115                 120
Trp Asn Ile Leu Lys Pro Leu Ile Ser His Lys Phe Gly Lys Lys
               125                 130                 135
Val Ile Tyr Phe Asn Tyr Leu Ser Glu Leu His Glu His Leu Lys
               140                 145                 150
Tyr Asp Gln Leu Val Ile Pro Pro Glu Val Leu Arg Tyr Asp Glu
               155                 160                 165
Lys Leu Gln Ser Leu His Glu Gly Arg Thr Pro Pro Pro Thr Lys
               170                 175                 180
```

Thr Pro Pro Pro Arg Pro Pro Leu Pro Thr Gln Gln Phe Gly Val
185 190 195
Ser Leu Gln Tyr Leu Lys Asp Lys Asn Gln Gly Glu Leu Ile Pro
200 205 210
Pro Val Leu Arg Phe Thr Val Thr Tyr Leu Arg Glu Lys Gly Leu
215 220 225
Arg Thr Glu Gly Leu Phe Arg Arg Ser Ala Ser Val Gln Thr Val
230 235 240
Arg Glu Ile Gln Arg Leu Tyr Asn Gln Gly Lys Pro Val Asn Phe
245 250 255
Asp Asp Tyr Gly Asp Ile His Ile Pro Ala Val Ile Leu Lys Thr
260 265 270
Phe Leu Arg Glu Leu Pro Gln Pro Leu Leu Thr Phe Gln Ala Tyr
275 280 285
Glu Gln Ile Leu Gly Ile Thr Cys Val Glu Ser Ser Leu Arg Val
290 295 300
Thr Gly Cys Arg Gln Ile Leu Arg Ser Leu Pro Glu His Asn Tyr
305 310 315
Val Val Leu Arg Tyr Leu Met Gly Phe Leu His Ala Val Ser Arg
320 325 330
Glu Ser Ile Phe Asn Lys Met Asn Ser Ser Asn Leu Ala Cys Val
335 340 345
Phe Gly Leu Asn Leu Ile Trp Pro Ser Gln Gly Val Ser Ser Leu
350 355 360
Ser Ala Leu Val Pro Leu Asn Met Phe Thr Glu Leu Leu Ile Glu
365 370 375
Tyr Tyr Glu Lys Ile Phe Ser Thr Pro Glu Ala Pro Gly Glu His
380 385 390
Gly Leu Ala Pro Trp Glu Gln Gly Ser Arg Ala Ala Pro Leu Gln
395 400 405
Glu Ala Val Pro Arg Thr Gln Ala Thr Gly Leu Thr Lys Pro Thr
410 415 420
Leu Pro Pro Ser Pro Leu Met Ala Ala Arg Arg Arg Leu
425 430

<210> 23
<211> 406
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5159072CD1

<400> 23
Met Ala Asp Gly Asn Glu Asp Leu Arg Ala Asp Asp Leu Pro Gly
1 5 10 15
Pro Ala Phe Glu Ser Tyr Glu Ser Met Glu Leu Ala Cys Pro Ala
20 25 30
Glu Arg Ser Gly His Val Ala Val Ser Asp Gly Arg His Met Phe
35 40 45
Val Trp Gly Gly Tyr Lys Ser Asn Gln Val Arg Gly Leu Tyr Asp
50 55 60
Phe Tyr Leu Pro Arg Glu Glu Leu Trp Ile Tyr Asn Met Glu Thr
65 70 75
Gly Arg Trp Lys Lys Ile Asn Thr Glu Gly Asp Val Pro Pro Ser
80 85 90
Met Ser Gly Ser Cys Ala Val Cys Val Asp Arg Val Leu Tyr Leu
95 100 105
Phe Gly Gly His His Ser Arg Gly Asn Thr Asn Lys Phe Tyr Met
110 115 120
Leu Asp Ser Arg Ser Thr Asp Arg Val Leu Gln Trp Glu Arg Ile
125 130 135
Asp Cys Gln Gly Ile Pro Pro Ser Ser Lys Asp Lys Leu Gly Val

```
                140                 145                 150
Trp Val Tyr Lys Asn Lys Leu Ile Phe Phe Gly Gly Tyr Gly Tyr
                155                 160                 165
Leu Pro Glu Asp Lys Val Leu Gly Thr Phe Glu Phe Asp Glu Thr
                170                 175                 180
Ser Phe Trp Asn Ser Ser His Pro Arg Gly Trp Asn Asp His Val
                185                 190                 195
His Ile Leu Asp Thr Glu Thr Phe Thr Trp Ser Gln Pro Ile Thr
                200                 205                 210
Thr Gly Lys Ala Pro Ser Pro Arg Ala Ala His Ala Cys Ala Thr
                215                 220                 225
Val Gly Asn Arg Gly Phe Val Phe Gly Gly Arg Tyr Arg Asp Ala
                230                 235                 240
Arg Met Asn Asp Leu His Tyr Leu Asn Leu Asp Thr Trp Glu Trp
                245                 250                 255
Asn Glu Leu Ile Pro Gln Gly Ile Cys Pro Val Gly Arg Ser Trp
                260                 265                 270
His Ser Leu Thr Pro Val Ser Ser Asp His Leu Phe Leu Phe Gly
                275                 280                 285
Gly Phe Thr Thr Asp Lys Gln Pro Leu Ser Asp Ala Trp Thr Tyr
                290                 295                 300
Cys Ile Ser Lys Asn Glu Trp Ile Gln Phe Asn His Pro Tyr Thr
                305                 310                 315
Glu Lys Pro Arg Leu Trp His Thr Ala Cys Ala Ser Asp Glu Gly
                320                 325                 330
Glu Val Ile Val Phe Gly Gly Cys Ala Asn Asn Leu Leu Val His
                335                 340                 345
His Arg Ala Ala His Ser Asn Glu Ile Leu Ile Phe Ser Val Gln
                350                 355                 360
Pro Lys Ser Leu Val Arg Leu Ser Leu Glu Ala Val Ile Cys Phe
                365                 370                 375
Lys Glu Met Leu Ala Asn Ser Trp Asn Cys Leu Pro Lys His Leu
                380                 385                 390
Leu His Ser Val Asn Gln Arg Phe Gly Ser Asn Asn Thr Ser Gly
                395                 400                 405
Ser

<210> 24
<211> 229
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5519057CD1

<400> 24
Met Ala Glu Glu Met Glu Ser Ser Leu Glu Ala Ser Phe Ser Ser
  1               5                  10                  15
Ser Gly Ala Val Ser Gly Ala Ser Gly Phe Leu Pro Pro Ala Arg
                 20                  25                  30
Ser Arg Ile Phe Lys Ile Ile Val Ile Gly Asp Ser Asn Val Gly
                 35                  40                  45
Lys Thr Cys Leu Thr Tyr Arg Phe Cys Ala Gly Arg Phe Pro Asp
                 50                  55                  60
Arg Thr Glu Ala Thr Ile Gly Val Asp Phe Arg Glu Arg Ala Val
                 65                  70                  75
Glu Ile Asp Gly Glu Arg Ile Lys Ile Gln Leu Trp Asp Thr Ala
                 80                  85                  90
Gly Gln Glu Arg Phe Arg Lys Ser Met Val Gln His Tyr Tyr Arg
                 95                 100                 105
Asn Val His Ala Val Val Phe Val Tyr Asp Met Thr Asn Met Ala
                110                 115                 120
```

```
Ser Phe His Ser Leu Pro Ser Trp Ile Glu Glu Cys Lys Gln His
                125                 130                 135
Leu Leu Ala Asn Asp Ile Pro Arg Ile Leu Val Gly Asn Lys Cys
                140                 145                 150
Asp Leu Arg Ser Ala Ile Gln Val Pro Thr Asp Leu Ala Gln Lys
                155                 160                 165
Phe Ala Asp Thr His Ser Met Pro Leu Phe Glu Thr Ser Ala Lys
                170                 175                 180
Asn Pro Asn Asp Asn Asp His Val Glu Ala Ile Phe Met Thr Leu
                185                 190                 195
Ala His Lys Leu Lys Cys His Lys Pro Leu Met Leu Ser Gln Pro
                200                 205                 210
Pro Asp Asn Gly Ile Ile Leu Lys Pro Glu Pro Lys Pro Ala Met
                215                 220                 225
Thr Cys Trp Cys

<210> 25
<211> 670
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 035379CD1

<400> 25
Met Ser Ser Gly Lys Ser Ala Arg Tyr Asn Arg Phe Ser Gly Gly
  1               5                  10                  15
Pro Ser Asn Leu Pro Thr Pro Asp Val Thr Thr Gly Thr Arg Met
                 20                  25                  30
Glu Thr Thr Phe Gly Pro Ala Phe Ser Ala Val Thr Thr Ile Thr
                 35                  40                  45
Lys Ala Asp Gly Thr Ser Thr Tyr Lys Gln His Cys Arg Thr Pro
                 50                  55                  60
Ser Ser Ser Ser Thr Leu Ala Tyr Ser Pro Arg Asp Glu Glu Asp
                 65                  70                  75
Ser Met Pro Pro Ile Ser Thr Pro Arg Arg Ser Asp Ser Ala Ile
                 80                  85                  90
Ser Val Arg Ser Leu His Ser Glu Ser Ser Met Ser Leu Arg Ser
                 95                 100                 105
Thr Phe Ser Leu Pro Glu Glu Glu Glu Glu Pro Glu Pro Leu Val
                110                 115                 120
Phe Ala Glu Gln Pro Ser Val Lys Leu Cys Cys Gln Leu Cys Cys
                125                 130                 135
Ser Val Phe Lys Asp Pro Val Ile Thr Thr Cys Gly His Thr Phe
                140                 145                 150
Cys Arg Arg Cys Ala Leu Lys Ser Glu Lys Cys Pro Val Asp Asn
                155                 160                 165
Val Lys Leu Thr Val Val Val Asn Asn Ile Ala Val Ala Glu Gln
                170                 175                 180
Ile Gly Glu Leu Phe Ile His Cys Arg His Gly Cys Arg Val Ala
                185                 190                 195
Gly Ser Gly Lys Pro Pro Ile Phe Glu Val Asp Pro Arg Gly Cys
                200                 205                 210
Pro Phe Thr Ile Lys Leu Ser Ala Arg Lys Asp His Glu Gly Ser
                215                 220                 225
Cys Asp Tyr Arg Pro Val Arg Cys Pro Asn Asn Pro Ser Cys Pro
                230                 235                 240
Pro Leu Leu Arg Met Asn Leu Glu Ala His Leu Lys Glu Cys Glu
                245                 250                 255
His Ile Lys Cys Pro His Ser Lys Tyr Gly Cys Thr Phe Ile Gly
                260                 265                 270
Asn Gln Asp Thr Tyr Glu Thr His Leu Glu Thr Cys Arg Phe Glu
```

```
                    275                 280                 285
Gly Leu Lys Glu Phe Leu Gln Gln Thr Asp Asp Arg Phe His Glu
                    290                 295                 300
Met His Val Ala Leu Ala Gln Lys Asp Gln Glu Ile Ala Phe Leu
                    305                 310                 315
Arg Ser Met Leu Gly Lys Leu Ser Glu Lys Ile Asp Gln Leu Glu
                    320                 325                 330
Lys Ser Leu Glu Leu Lys Phe Asp Val Leu Asp Glu Asn Gln Ser
                    335                 340                 345
Lys Leu Ser Glu Asp Leu Met Glu Phe Arg Arg Asp Ala Ser Met
                    350                 355                 360
Leu Asn Asp Glu Leu Ser His Ile Asn Ala Arg Leu Asn Met Gly
                    365                 370                 375
Ile Leu Gly Ser Tyr Asp Pro Gln Gln Ile Phe Lys Cys Lys Gly
                    380                 385                 390
Thr Phe Val Gly His Gln Gly Pro Val Trp Cys Leu Cys Val Tyr
                    395                 400                 405
Ser Met Gly Asp Leu Leu Phe Ser Gly Ser Ser Asp Lys Thr Ile
                    410                 415                 420
Lys Val Trp Asp Thr Cys Thr Thr Tyr Lys Cys Gln Lys Thr Leu
                    425                 430                 435
Glu Gly His Asp Gly Ile Val Leu Ala Leu Cys Ile Gln Gly Cys
                    440                 445                 450
Lys Leu Tyr Ser Gly Ser Ala Asp Cys Thr Ile Ile Val Trp Asp
                    455                 460                 465
Ile Gln Asn Leu Gln Lys Val Asn Thr Ile Arg Ala His Asp Asn
                    470                 475                 480
Pro Val Cys Thr Leu Val Ser Ser His Asn Val Leu Phe Ser Gly
                    485                 490                 495
Ser Leu Lys Ala Ile Lys Val Trp Asp Ile Val Gly Thr Glu Leu
                    500                 505                 510
Lys Leu Lys Lys Glu Leu Thr Gly Leu Asn His Trp Val Arg Ala
                    515                 520                 525
Leu Val Ala Ala Gln Ser Tyr Leu Tyr Ser Gly Ser Tyr Gln Thr
                    530                 535                 540
Ile Lys Ile Trp Asp Ile Arg Thr Leu Asp Cys Ile His Val Leu
                    545                 550                 555
Gln Thr Ser Gly Gly Ser Val Tyr Ser Ile Ala Val Thr Asn His
                    560                 565                 570
His Ile Val Cys Gly Thr Tyr Glu Asn Leu Ile His Val Trp Asp
                    575                 580                 585
Ile Glu Ser Lys Glu Gln Val Arg Thr Leu Thr Gly His Val Gly
                    590                 595                 600
Thr Val Tyr Ala Leu Ala Val Ile Ser Thr Pro Asp Gln Thr Lys
                    605                 610                 615
Val Phe Ser Ala Ser Tyr Asp Arg Ser Leu Arg Val Trp Ser Met
                    620                 625                 630
Asp Asn Met Ile Cys Thr Gln Thr Leu Leu Arg His Gln Ser Ser
                    635                 640                 645
Val Thr Ala Leu Ala Val Ser Arg Gly Arg Leu Phe Ser Gly Ala
                    650                 655                 660
Val Asp Ser Thr Val Lys Val Trp Thr Cys
                    665                 670
```

<210> 26
<211> 445
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 275354CD1

<400> 26

```
Met Lys Val Lys Met Leu Ser Arg Asn Pro Asp Asn Tyr Val Arg
 1               5                  10                  15
Glu Thr Lys Leu Asp Leu Gln Arg Val Pro Arg Asn Tyr Asp Pro
                20                  25                  30
Ala Leu His Pro Phe Glu Val Pro Arg Glu Tyr Val Arg Ala Leu
                35                  40                  45
Asn Ala Thr Lys Leu Glu Arg Val Phe Ala Lys Pro Phe Leu Ala
                50                  55                  60
Ser Leu Asp Gly His Arg Asp Gly Val Asn Cys Leu Ala Lys His
                65                  70                  75
Pro Glu Lys Leu Ala Thr Val Leu Ser Gly Ala Cys Asp Gly Glu
                80                  85                  90
Val Arg Ile Trp Asn Leu Thr Gln Arg Asn Cys Ile Arg Thr Ile
                95                  100                 105
Gln Ala His Glu Gly Phe Val Arg Gly Ile Cys Thr Arg Phe Cys
                110                 115                 120
Gly Thr Ser Phe Phe Thr Val Gly Asp Asp Lys Thr Val Lys Gln
                125                 130                 135
Trp Lys Met Asp Gly Pro Gly Tyr Gly Asp Glu Glu Glu Pro Leu
                140                 145                 150
His Thr Ile Leu Gly Lys Thr Val Tyr Thr Gly Ile Asp His His
                155                 160                 165
Trp Lys Glu Ala Val Phe Ala Thr Cys Gly Gln Gln Val Asp Ile
                170                 175                 180
Trp Asp Glu Gln Arg Thr Asn Pro Ile Cys Ser Met Thr Trp Gly
                185                 190                 195
Phe Asp Ser Ile Ser Ser Val Lys Phe Asn Pro Ile Glu Thr Phe
                200                 205                 210
Leu Leu Gly Ser Cys Ala Ser Asp Arg Asn Ile Val Leu Tyr Asp
                215                 220                 225
Met Arg Gln Ala Thr Pro Leu Lys Lys Val Ile Leu Asp Met Arg
                230                 235                 240
Thr Asn Thr Ile Cys Trp Asn Pro Met Glu Ala Phe Ile Phe Thr
                245                 250                 255
Ala Ala Asn Glu Asp Tyr Asn Leu Tyr Thr Phe Asp Met Arg Ala
                260                 265                 270
Leu Asp Thr Pro Val Met Val His Met Asp His Val Ser Ala Val
                275                 280                 285
Leu Asp Val Asp Tyr Ser Pro Thr Gly Lys Glu Phe Val Ser Ala
                290                 295                 300
Ser Phe Asp Lys Ser Ile Arg Ile Phe Pro Val Asp Lys Ser Arg
                305                 310                 315
Ser Arg Glu Val Tyr His Thr Lys Arg Met Gln His Val Ile Cys
                320                 325                 330
Val Lys Trp Thr Ser Asp Ser Lys Tyr Ile Met Cys Gly Ser Asp
                335                 340                 345
Glu Met Asn Ile Arg Leu Trp Lys Ala Asn Ala Ser Glu Lys Leu
                350                 355                 360
Gly Val Leu Thr Ser Arg Glu Lys Ala Ala Lys Asp Tyr Asn Gln
                365                 370                 375
Lys Leu Lys Glu Lys Phe Gln His Tyr Pro His Ile Lys Arg Ile
                380                 385                 390
Ala Arg His Arg His Leu Pro Lys Ser Ile Tyr Ser Gln Ile Gln
                395                 400                 405
Glu Gln Arg Ile Met Lys Glu Ala Arg Arg Arg Lys Glu Val Asn
                410                 415                 420
Arg Ile Lys His Ser Lys Pro Gly Ser Val Pro Leu Val Ser Glu
                425                 430                 435
Lys Lys Lys His Val Val Ala Val Val Lys
                440                 445
```

<210> 27
<211> 236
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 311658CD1

<400> 27
```
Met Ser Asp Leu Leu Ser Pro Leu Leu Tyr Val Met Glu Asn Glu
 1               5                  10                  15
Val Asp Ala Phe Trp Cys Phe Ala Ser Tyr Met Asp Gln Met His
                20                  25                  30
Gln Asn Phe Glu Glu Gln Met Gln Gly Met Lys Thr Gln Leu Ile
                35                  40                  45
Gln Leu Ser Thr Leu Leu Arg Leu Leu Asp Ser Gly Phe Cys Ser
                50                  55                  60
Tyr Leu Glu Ser Gln Asp Ser Gly Tyr Leu Tyr Phe Cys Phe Arg
                65                  70                  75
Trp Leu Leu Ile Arg Phe Lys Arg Glu Phe Ser Phe Leu Asp Ile
                80                  85                  90
Leu Arg Leu Trp Glu Val Met Trp Thr Glu Leu Pro Cys Thr Asn
                95                 100                 105
Phe His Leu Leu Leu Cys Cys Ala Ile Leu Glu Ser Glu Lys Gln
               110                 115                 120
Gln Ile Met Glu Lys His Tyr Gly Phe Asn Glu Ile Leu Lys His
               125                 130                 135
Ile Asn Glu Leu Ser Met Lys Ile Asp Val Glu Asp Ile Leu Cys
               140                 145                 150
Lys Ala Glu Ala Ile Ser Leu Gln Met Val Lys Cys Lys Glu Leu
               155                 160                 165
Pro Gln Ala Val Cys Glu Ile Leu Gly Leu Gln Gly Ser Glu Val
               170                 175                 180
Thr Thr Pro Asp Ser Asp Val Gly Glu Asp Glu Asn Val Val Met
               185                 190                 195
Thr Pro Cys Pro Thr Ser Ala Phe Gln Ser Asn Ala Leu Pro Thr
               200                 205                 210
Leu Ser Ala Ser Gly Ala Arg Asn Asp Ser Pro Thr Gln Ile Pro
               215                 220                 225
Val Ser Ser Asp Val Cys Arg Leu Thr Pro Ala
               230                 235
```
<210> 28
<211> 498
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1251632CD1

<400> 28
```
Met Gln Glu Ser Gly Cys Arg Leu Glu His Pro Ser Ala Thr Lys
 1               5                  10                  15
Phe Arg Asn His Val Met Glu Gly Asp Trp Asp Lys Ala Glu Asn
                20                  25                  30
Asp Leu Asn Glu Leu Lys Pro Leu Val His Ser Pro His Ala Ile
                35                  40                  45
Val Arg Met Lys Phe Leu Leu Leu Gln Gln Lys Tyr Leu Glu Tyr
                50                  55                  60
Leu Glu Asp Gly Lys Val Leu Glu Ala Leu Gln Val Leu Arg Cys
                65                  70                  75
Glu Leu Thr Pro Leu Lys Tyr Asn Thr Glu Arg Ile His Val Leu
                80                  85                  90
Ser Gly Tyr Leu Met Cys Ser His Ala Glu Asp Leu Arg Ala Lys
                95                 100                 105
```

```
Ala Glu Trp Glu Gly Lys Gly Thr Ala Ser Arg Ser Lys Leu Leu
                110                 115                 120
Asp Lys Leu Gln Thr Tyr Leu Pro Pro Ser Val Met Leu Pro Pro
                125                 130                 135
Arg Arg Leu Gln Thr Leu Leu Arg Gln Ala Val Glu Leu Gln Arg
                140                 145                 150
Asp Arg Cys Leu Tyr His Asn Thr Lys Leu Asp Asn Asn Leu Asp
                155                 160                 165
Ser Val Ser Leu Leu Ile Asp His Val Cys Ser Arg Arg Gln Phe
                170                 175                 180
Pro Cys Tyr Thr Gln Gln Ile Leu Thr Glu His Cys Asn Glu Val
                185                 190                 195
Trp Phe Cys Lys Phe Ser Asn Asp Gly Thr Lys Leu Ala Thr Gly
                200                 205                 210
Ser Lys Asp Thr Thr Val Ile Ile Trp Gln Val Asp Pro Asp Thr
                215                 220                 225
His Leu Leu Lys Leu Leu Lys Thr Leu Glu Gly His Ala Tyr Gly
                230                 235                 240
Val Ser Tyr Ile Ala Trp Ser Pro Asp Asp Asn Tyr Leu Val Ala
                245                 250                 255
Cys Gly Pro Asp Asp Cys Ser Glu Leu Trp Leu Trp Asn Val Gln
                260                 265                 270
Thr Gly Glu Leu Arg Thr Lys Met Ser Gln Ser His Glu Asp Ser
                275                 280                 285
Leu Thr Ser Val Ala Trp Asn Pro Asp Gly Lys Arg Phe Val Thr
                290                 295                 300
Gly Gly Gln Arg Gly Gln Phe Tyr Gln Cys Asp Leu Asp Gly Asn
                305                 310                 315
Leu Leu Asp Ser Trp Glu Gly Val Arg Val Gln Cys Leu Trp Cys
                320                 325                 330
Leu Ser Asp Gly Lys Thr Val Leu Ala Ser Asp Thr His Gln Arg
                335                 340                 345
Ile Arg Gly Tyr Asn Phe Glu Asp Leu Thr Asp Arg Asn Ile Val
                350                 355                 360
Gln Glu Asp His Pro Ile Met Ser Phe Thr Ile Ser Lys Asn Gly
                365                 370                 375
Arg Leu Ala Leu Leu Asn Val Ala Thr Gln Gly Val His Leu Trp
                380                 385                 390
Asp Leu Gln Asp Arg Val Leu Val Arg Lys Tyr Gln Gly Val Thr
                395                 400                 405
Gln Gly Phe Tyr Thr Ile His Ser Cys Phe Gly Gly His Asn Glu
                410                 415                 420
Asp Phe Ile Ala Ser Gly Ser Glu Asp His Lys Val Tyr Ile Trp
                425                 430                 435
His Lys Arg Ser Glu Leu Pro Ile Ala Glu Leu Thr Gly His Thr
                440                 445                 450
Arg Thr Val Asn Cys Val Ser Trp Asn Pro Gln Ile Pro Ser Met
                455                 460                 465
Met Ala Ser Ala Ser Asp Asp Gly Thr Val Arg Ile Trp Gly Pro
                470                 475                 480
Ala Pro Phe Ile Asp His Gln Asn Ile Glu Glu Glu Cys Ser Ser
                485                 490                 495
Met Asp Ser
```

<210> 29
<211> 334
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1331955CD1

<400> 29
Met Ala Thr Glu Glu Lys Lys Pro Glu Thr Glu Ala Ala Arg Ala
1 5 10 15
Gln Pro Thr Pro Ser Ser Ser Ala Thr Gln Ser Lys Pro Thr Pro
20 25 30
Val Lys Pro Asn Tyr Ala Leu Lys Phe Thr Leu Ala Gly His Thr
35 40 45
Lys Ala Val Ser Ser Val Lys Phe Ser Pro Asn Gly Glu Trp Leu
50 55 60
Ala Ser Ser Ser Ala Asp Lys Leu Ile Lys Ile Trp Gly Ala Tyr
65 70 75
Asp Gly Lys Phe Glu Lys Thr Ile Ser Gly His Lys Leu Gly Ile
80 85 90
Ser Asp Val Ala Trp Ser Ser Asp Ser Asn Leu Leu Val Ser Ala
95 100 105
Ser Asp Asp Lys Thr Leu Lys Ile Trp Asp Val Ser Ser Gly Lys
110 115 120
Cys Leu Lys Thr Leu Lys Gly His Ser Asn Tyr Val Phe Cys Cys
125 130 135
Asn Phe Asn Pro Gln Ser Asn Leu Ile Val Ser Gly Ser Phe Asp
140 145 150
Glu Ser Val Arg Ile Trp Asp Val Lys Thr Gly Lys Cys Leu Lys
155 160 165
Thr Leu Pro Ala His Ser Asp Pro Val Ser Ala Val His Phe Asn
170 175 180
Arg Asp Gly Ser Leu Ile Val Ser Ser Ser Tyr Asp Gly Leu Cys
185 190 195
Arg Ile Trp Asp Thr Ala Ser Gly Gln Cys Leu Lys Thr Leu Ile
200 205 210
Asp Asp Asp Asn Pro Pro Val Ser Phe Val Lys Phe Ser Pro Asn
215 220 225
Gly Lys Tyr Ile Leu Ala Ala Thr Leu Asp Asn Thr Leu Lys Leu
230 235 240
Trp Asp Tyr Ser Lys Gly Lys Cys Leu Lys Thr Tyr Thr Gly His
245 250 255
Lys Asn Glu Lys Tyr Cys Ile Phe Ala Asn Phe Ser Val Thr Gly
260 265 270
Gly Lys Trp Ile Val Ser Gly Ser Glu Asp Asn Leu Val Tyr Ile
275 280 285
Trp Asn Leu Gln Thr Lys Glu Ile Val Gln Lys Leu Gln Gly His
290 295 300
Thr Asp Val Val Ile Ser Thr Ala Cys His Pro Thr Glu Asn Ile
305 310 315
Ile Ala Ser Ala Ala Leu Glu Asn Asp Lys Thr Ile Lys Leu Trp
320 325 330
Lys Ser Asp Cys

<210> 30
<211> 292
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1412614CD1

<400> 30
Met Met Ala Phe Ala Pro Pro Lys Asn Thr Asp Gly Pro Lys Met
1 5 10 15
Gln Thr Lys Met Ser Thr Trp Thr Pro Leu Asn His Gln Leu Leu
20 25 30
Asn Asp Arg Val Phe Glu Glu Arg Arg Ala Leu Leu Gly Lys Trp
35 40 45

```
Phe Asp Lys Trp Thr Asp Ser Gln Arg Arg Arg Ile Leu Thr Gly
                 50                  55                  60
Leu Leu Glu Arg Cys Ser Leu Ser Gln Gln Lys Phe Cys Cys Arg
                 65                  70                  75
Lys Leu Gln Glu Lys Ile Pro Ala Glu Ala Leu Asp Phe Thr Thr
                 80                  85                  90
Lys Leu Pro Arg Val Leu Ser Leu Tyr Ile Phe Ser Phe Leu Asp
                 95                 100                 105
Pro Arg Ser Leu Cys Arg Cys Ala Gln Val Cys Trp His Trp Lys
                110                 115                 120
Asn Leu Ala Glu Leu Asp Gln Leu Trp Met Leu Lys Cys Leu Arg
                125                 130                 135
Phe Asn Trp Tyr Ile Asn Phe Ser Pro Thr Pro Phe Glu Gln Gly
                140                 145                 150
Ile Trp Lys Lys His Tyr Ile Gln Met Val Lys Glu Leu His Ile
                155                 160                 165
Thr Lys Pro Lys Thr Pro Pro Lys Asp Gly Phe Val Ile Ala Asp
                170                 175                 180
Val Gln Leu Val Thr Ser Asn Ser Pro Glu Glu Lys Gln Ser Pro
                185                 190                 195
Leu Ser Ala Phe Arg Ser Ser Ser Ser Leu Arg Lys Lys Asn Asn
                200                 205                 210
Ser Gly Glu Lys Ala Leu Pro Pro Trp Arg Ser Ser Asp Lys His
                215                 220                 225
Pro Thr Asp Ile Ile Arg Phe Asn Tyr Leu Asp Asn Arg Asp Pro
                230                 235                 240
Met Glu Thr Val Gln Gln Gly Arg Arg Lys Arg Asn Gln Ile Thr
                245                 250                 255
Pro Asp Phe Ser Arg Gln Ser His Asp Lys Lys Asn Lys Leu Gln
                260                 265                 270
Asp Arg Thr Arg Leu Arg Lys Ala Gln Ser Met Met Ser Arg Arg
                275                 280                 285
Asn Pro Phe Pro Leu Cys Pro
                290

<210> 31
<211> 588
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1750781CD1

<400> 31
Met Ser Ser Gly Leu Arg Ala Ala Asp Phe Pro Arg Trp Lys Arg
  1               5                  10                  15
His Ile Ser Glu Gln Leu Arg Arg Arg Asp Arg Leu Gln Arg Gln
                 20                  25                  30
Ala Phe Glu Glu Ile Ile Leu Gln Tyr Asn Lys Leu Leu Glu Lys
                 35                  40                  45
Ser Asp Leu His Ser Val Leu Ala Gln Lys Leu Gln Ala Glu Lys
                 50                  55                  60
His Asp Val Pro Asn Arg His Glu Ile Ser Pro Gly His Asp Gly
                 65                  70                  75
Thr Trp Asn Asp Asn Gln Leu Gln Glu Met Ala Gln Leu Arg Ile
                 80                  85                  90
Lys His Gln Glu Glu Leu Thr Glu Leu His Lys Lys Arg Gly Glu
                 95                 100                 105
Leu Ala Gln Leu Val Ile Asp Leu Asn Asn Gln Met Gln Arg Lys
                110                 115                 120
Asp Arg Glu Met Gln Met Asn Glu Ala Lys Ile Ala Glu Cys Leu
                125                 130                 135
Gln Thr Ile Ser Asp Leu Glu Thr Glu Cys Leu Asp Leu Arg Thr
```

```
                140                           145                           150
Lys Leu Cys Asp Leu Glu Arg Ala Asn Gln Thr Leu Lys Asp Glu
                155                           160                           165
Tyr Asp Ala Leu Gln Ile Thr Phe Thr Ala Leu Glu Gly Lys Leu
                170                           175                           180
Arg Lys Thr Thr Glu Glu Asn Gln Glu Leu Val Thr Arg Trp Met
                185                           190                           195
Ala Glu Lys Ala Gln Glu Ala Asn Arg Leu Asn Ala Glu Asn Glu
                200                           205                           210
Lys Asp Ser Arg Arg Arg Gln Ala Arg Leu Gln Lys Glu Leu Ala
                215                           220                           225
Glu Ala Ala Lys Glu Pro Leu Pro Val Glu Gln Asp Asp Asp Ile
                230                           235                           240
Glu Val Ile Val Asp Glu Thr Ser Asp His Thr Glu Glu Thr Ser
                245                           250                           255
Pro Val Arg Ala Ile Ser Arg Ala Ala Thr Arg Arg Ser Val Ser
                260                           265                           270
Ser Phe Pro Val Pro Gln Asp Asn Val Asp Thr His Pro Gly Ser
                275                           280                           285
Gly Lys Glu Val Arg Val Pro Ala Thr Ala Leu Cys Val Phe Asp
                290                           295                           300
Ala His Asp Gly Glu Val Asn Ala Val Gln Phe Ser Pro Gly Ser
                305                           310                           315
Arg Leu Leu Ala Thr Gly Gly Met Asp Arg Arg Val Lys Leu Trp
                320                           325                           330
Glu Val Phe Gly Glu Lys Cys Glu Phe Lys Gly Ser Leu Ser Gly
                335                           340                           345
Ser Asn Ala Gly Ile Thr Ser Ile Glu Phe Asp Ser Ala Gly Ser
                350                           355                           360
Tyr Leu Leu Ala Ala Ser Asn Asp Phe Ala Ser Arg Ile Trp Thr
                365                           370                           375
Val Asp Asp Tyr Arg Leu Arg His Thr Leu Thr Gly His Ser Gly
                380                           385                           390
Lys Val Leu Ser Ala Lys Phe Leu Leu Asp Asn Ala Arg Ile Val
                395                           400                           405
Ser Gly Ser His Asp Arg Thr Leu Lys Leu Trp Asp Leu Arg Ser
                410                           415                           420
Lys Val Cys Ile Lys Thr Val Phe Ala Gly Ser Ser Cys Asn Asp
                425                           430                           435
Ile Val Cys Thr Glu Gln Cys Val Met Ser Gly His Phe Asp Lys
                440                           445                           450
Lys Ile Arg Phe Trp Asp Ile Arg Ser Glu Ser Ile Val Arg Glu
                455                           460                           465
Met Glu Leu Leu Gly Lys Ile Thr Ala Leu Asp Leu Asn Pro Glu
                470                           475                           480
Arg Thr Glu Leu Leu Ser Cys Ser Arg Asp Asp Leu Leu Lys Val
                485                           490                           495
Ile Asp Leu Arg Thr Asn Ala Ile Lys Gln Thr Phe Ser Ala Pro
                500                           505                           510
Gly Phe Lys Cys Gly Ser Asp Trp Thr Arg Val Val Phe Ser Pro
                515                           520                           525
Asp Gly Ser Tyr Val Ala Ala Gly Ser Ala Glu Gly Ser Leu Tyr
                530                           535                           540
Ile Trp Ser Val Leu Thr Gly Lys Val Glu Lys Val Leu Ser Lys
                545                           550                           555
Gln His Ser Ser Ser Ile Asn Ala Val Ala Trp Ser Pro Ser Gly
                560                           565                           570
Ser His Val Val Ser Val Asp Lys Gly Cys Lys Ala Val Leu Trp
                575                           580                           585
Ala Gln Tyr
```

<210> 32
<211> 326

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1821658CD1

<400> 32
Met Lys Gln Asp Ala Ser Arg Asn Ala Ala Tyr Thr Val Asp Cys
1 5 10 15
Glu Asp Tyr Val His Val Val Glu Phe Asn Pro Phe Glu Asn Gly
20 25 30
Asp Ser Gly Asn Leu Ile Ala Tyr Gly Gly Asn Asn Tyr Val Val
35 40 45
Ile Gly Thr Cys Thr Phe Gln Glu Glu Glu Ala Asp Val Glu Gly
50 55 60
Ile Gln Tyr Lys Thr Leu Arg Thr Phe His His Gly Val Arg Val
65 70 75
Asp Gly Ile Ala Trp Ser Pro Glu Thr Arg Leu Asp Ser Leu Pro
80 85 90
Pro Val Ile Lys Phe Cys Thr Ser Ala Ala Asp Met Lys Ile Arg
95 100 105
Leu Phe Thr Ser Asp Leu Gln Asp Lys Asn Glu Tyr Lys Val Leu
110 115 120
Glu Gly His Thr Asp Phe Ile Asn Gly Leu Val Phe Asp Pro Lys
125 130 135
Glu Gly Gln Glu Ile Ala Ser Val Ser Asp Asp His Thr Cys Arg
140 145 150
Ile Trp Asn Leu Glu Gly Val Gln Thr Ala His Phe Val Leu His
155 160 165
Ser Pro Gly Met Ser Val Cys Trp His Pro Glu Glu Thr Phe Lys
170 175 180
Leu Met Val Ala Glu Lys Asn Gly Thr Ile Arg Phe Tyr Asp Leu
185 190 195
Leu Ala Gln Gln Ala Ile Leu Ser Leu Glu Ser Glu Gln Val Pro
200 205 210
Leu Met Ser Ala His Trp Cys Leu Lys Asn Thr Phe Lys Val Gly
215 220 225
Ala Val Ala Gly Asn Asp Trp Leu Ile Trp Asp Ile Thr Arg Ser
230 235 240
Ser Tyr Pro Gln Asn Lys Arg Pro Val His Met Asp Arg Ala Cys
245 250 255
Leu Phe Arg Trp Ser Thr Ile Ser Glu Asn Leu Phe Ala Thr Thr
260 265 270
Gly Tyr Pro Gly Lys Met Ala Ser Gln Phe Gln Ile His His Leu
275 280 285
Gly His Pro Gln Pro Ile Leu Met Gly Ser Val Ala Val Gly Ser
290 295 300
Gly Leu Ser Trp His Arg Thr Leu Pro Leu Cys Val Ile Gly Gly
305 310 315
Asp His Lys Leu Leu Phe Trp Val Thr Glu Val
320 325

<210> 33
<211> 453
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1872574CD1

<400> 33
Met Ala Arg Lys Val Val Ser Arg Lys Arg Lys Ala Pro Ala Ser

```
1               5                   10                  15
Pro Gly Ala Gly Ser Asp Ala Gln Gly Pro Gln Phe Gly Trp Asp
                20                  25                  30
His Ser Leu His Lys Arg Lys Arg Leu Pro Pro Val Lys Arg Ser
                35                  40                  45
Leu Val Tyr Tyr Leu Lys Asn Arg Glu Val Arg Leu Gln Asn Glu
                50                  55                  60
Thr Ser Tyr Ser Arg Val Leu His Gly Tyr Ala Ala Gln Gln Leu
                65                  70                  75
Pro Ser Leu Leu Lys Glu Arg Glu Phe His Leu Gly Thr Leu Asn
                80                  85                  90
Lys Val Phe Ala Ser Gln Trp Leu Asn His Arg Gln Val Val Cys
                95                  100                 105
Gly Thr Lys Cys Asn Thr Leu Phe Val Val Asp Val Gln Thr Ser
                110                 115                 120
Gln Ile Thr Lys Ile Pro Ile Leu Lys Asp Arg Glu Pro Gly Gly
                125                 130                 135
Val Thr Gln Gln Gly Cys Gly Ile His Ala Ile Glu Leu Asn Pro
                140                 145                 150
Ser Arg Thr Leu Leu Ala Thr Gly Gly Asp Asn Pro Asn Ser Leu
                155                 160                 165
Ala Ile Tyr Arg Leu Pro Thr Leu Asp Pro Val Cys Val Gly Asp
                170                 175                 180
Asp Gly His Lys Asp Trp Ile Phe Ser Ile Ala Trp Ile Ser Asp
                185                 190                 195
Thr Met Ala Val Ser Gly Ser Arg Asp Gly Ser Met Gly Leu Trp
                200                 205                 210
Glu Val Thr Asp Asp Val Leu Thr Lys Ser Asp Ala Arg His Asn
                215                 220                 225
Val Ser Arg Val Pro Val Tyr Ala His Ile Thr His Lys Ala Leu
                230                 235                 240
Lys Asp Ile Pro Lys Glu Asp Thr Asn Pro Asp Asn Cys Lys Val
                245                 250                 255
Arg Ala Leu Ala Phe Asn Asn Lys Asn Lys Glu Leu Gly Ala Val
                260                 265                 270
Ser Leu Asp Gly Tyr Phe His Leu Trp Lys Ala Glu Asn Thr Leu
                275                 280                 285
Ser Lys Leu Leu Ser Thr Lys Leu Pro Tyr Cys Arg Glu Asn Val
                290                 295                 300
Cys Leu Ala Tyr Gly Ser Glu Trp Ser Val Tyr Ala Val Gly Ser
                305                 310                 315
Gln Ala His Val Ser Phe Leu Asp Pro Arg Gln Pro Ser Tyr Asn
                320                 325                 330
Val Lys Ser Val Cys Ser Arg Glu Arg Gly Ser Gly Ile Arg Ser
                335                 340                 345
Val Ser Phe Tyr Glu His Ile Ile Thr Val Gly Thr Gly Gln Gly
                350                 355                 360
Ser Leu Leu Phe Tyr Asp Ile Arg Ala Gln Arg Phe Leu Glu Glu
                365                 370                 375
Arg Leu Ser Ala Cys Tyr Gly Ser Lys Pro Arg Leu Ala Gly Glu
                380                 385                 390
Asn Leu Lys Leu Thr Thr Gly Lys Gly Trp Leu Asn His Asp Glu
                395                 400                 405
Thr Trp Arg Asn Tyr Phe Ser Asp Ile Asp Phe Phe Pro Asn Ala
                410                 415                 420
Val Tyr Thr His Cys Tyr Asp Ser Ser Gly Thr Lys Leu Phe Val
                425                 430                 435
Ala Gly Gly Pro Leu Pro Ser Gly Leu His Gly Asn Tyr Ala Gly
                440                 445                 450
Leu Trp Ser
```

<210> 34
<211> 161

<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2590967CD1

<400> 34
Met Ala Thr Glu Gly Gly Gly Lys Glu Met Asn Glu Ile Lys Thr
1 5 10 15
Gln Phe Thr Thr Arg Glu Gly Leu Tyr Lys Leu Leu Pro His Ser
20 25 30
Glu Tyr Ser Arg Pro Asn Arg Val Pro Phe Asn Ser Gln Gly Ser
35 40 45
Asn Pro Val Arg Val Ser Phe Val Asn Leu Asn Asp Gln Ser Gly
50 55 60
Asn Gly Asp Arg Leu Cys Phe Asn Val Gly Arg Glu Leu Tyr Phe
65 70 75
Tyr Ile Tyr Lys Gly Val Arg Lys Ala Ala Asp Leu Ser Lys Pro
80 85 90
Ile Asp Lys Arg Ile Tyr Lys Gly Thr Gln Pro Thr Cys His Asp
95 100 105
Phe Asn His Leu Thr Ala Thr Ala Glu Ser Val Ser Leu Leu Val
110 115 120
Gly Phe Ser Ala Gly Gln Val Gln Leu Ile Asp Pro Ile Lys Lys
125 130 135
Glu Thr Ser Lys Leu Phe Asn Glu Glu Gly Ser Leu Ser Ser Pro
140 145 150
Ser Gln Ala Ser Ser Pro Gly Gly Thr Val Val
155 160

<210> 35
<211> 684
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2824491CD1

<400> 35
Met Ala Arg His Arg Asn Val Arg Gly Tyr Asn Tyr Asp Glu Asp
1 5 10 15
Phe Glu Asp Asp Asp Leu Tyr Gly Gln Ser Val Glu Asp Asp Tyr
20 25 30
Cys Ile Ser Pro Ser Thr Ala Ala Gln Phe Ile Tyr Ser Arg Arg
35 40 45
Asp Lys Pro Ser Val Glu Pro Val Glu Glu Tyr Asp Tyr Glu Asp
50 55 60
Leu Lys Glu Ser Ser Asn Ser Val Ser Asn His Gln Leu Ser Gly
65 70 75
Phe Asp Gln Ala Arg Leu Tyr Ser Cys Leu Asp His Met Arg Glu
80 85 90
Val Leu Gly Asp Ala Val Pro Asp Glu Ile Leu Ile Glu Ala Val
95 100 105
Leu Lys Asn Lys Phe Asp Val Gln Lys Ala Leu Ser Gly Val Leu
110 115 120
Glu Gln Asp Arg Val Gln Ser Leu Lys Asp Lys Asn Glu Ala Thr
125 130 135
Val Ser Thr Gly Lys Ile Ala Lys Gly Lys Pro Val Asp Ser Gln
140 145 150
Thr Ser Arg Ser Glu Ser Glu Ile Val Pro Lys Val Ala Lys Met
155 160 165
Thr Val Ser Gly Lys Lys Gln Thr Met Gly Phe Glu Val Pro Gly

```
                    170                      175                      180
Val Ser Ser Glu Glu Asn Gly His Ser Phe His Thr Pro Gln Lys
                    185                      190                      195
Gly Pro Pro Ile Glu Asp Ala Ile Ala Ser Ser Asp Val Leu Glu
                    200                      205                      210
Thr Ala Ser Lys Ser Ala Asn Pro Pro His Thr Ile Gln Ala Ser
                    215                      220                      225
Glu Glu Gln Ser Ser Thr Pro Ala Pro Val Lys Lys Ser Gly Lys
                    230                      235                      240
Leu Arg Gln Gln Ile Asp Val Lys Ala Glu Leu Glu Lys Arg Gln
                    245                      250                      255
Gly Gly Lys Gln Leu Leu Asn Leu Val Val Ile Gly His Val Asp
                    260                      265                      270
Ala Gly Lys Ser Thr Leu Met Gly His Met Leu Tyr Leu Leu Gly
                    275                      280                      285
Asn Ile Asn Lys Arg Thr Met His Lys Tyr Glu Gln Glu Ser Lys
                    290                      295                      300
Lys Ala Gly Lys Ala Ser Phe Ala Tyr Ala Trp Val Leu Asp Glu
                    305                      310                      315
Thr Gly Glu Glu Arg Glu Arg Gly Val Thr Met Asp Val Gly Met
                    320                      325                      330
Thr Lys Phe Glu Thr Thr Thr Lys Val Ile Thr Leu Met Asp Ala
                    335                      340                      345
Pro Gly His Lys Asp Phe Ile Pro Asn Met Ile Thr Gly Ala Ala
                    350                      355                      360
Gln Ala Asp Val Ala Val Leu Val Val Asp Ala Ser Arg Gly Glu
                    365                      370                      375
Phe Glu Ala Gly Phe Glu Thr Gly Gly Gln Thr Arg Glu His Gly
                    380                      385                      390
Leu Leu Val Arg Ser Leu Gly Val Thr Gln Leu Ala Val Ala Val
                    395                      400                      405
Asn Lys Met Asp Gln Val Asn Trp Gln Gln Glu Arg Phe Gln Glu
                    410                      415                      420
Ile Thr Gly Lys Leu Gly His Phe Leu Lys Gln Ala Gly Phe Lys
                    425                      430                      435
Glu Ser Asp Val Gly Phe Ile Pro Thr Ser Gly Leu Ser Gly Glu
                    440                      445                      450
Asn Leu Ile Thr Arg Ser Gln Ser Ser Glu Leu Thr Lys Trp Tyr
                    455                      460                      465
Lys Gly Leu Cys Leu Leu Glu Gln Ile Asp Ser Phe Lys Pro Pro
                    470                      475                      480
Gln Arg Ser Ile Asp Lys Pro Phe Arg Leu Cys Val Ser Asp Val
                    485                      490                      495
Phe Lys Asp Gln Gly Ser Gly Phe Cys Ile Thr Gly Lys Ile Glu
                    500                      505                      510
Ala Gly Tyr Ile Gln Thr Gly Asp Arg Leu Leu Ala Met Pro Pro
                    515                      520                      525
Asn Glu Thr Cys Thr Val Lys Gly Ile Thr Leu His Asp Glu Pro
                    530                      535                      540
Val Asp Trp Ala Ala Ala Gly Asp His Val Ser Leu Thr Leu Val
                    545                      550                      555
Gly Met Asp Ile Ile Lys Ile Asn Val Gly Cys Ile Phe Cys Gly
                    560                      565                      570
Pro Lys Val Pro Ile Lys Ala Cys Thr Arg Phe Arg Ala Arg Ile
                    575                      580                      585
Leu Ile Phe Asn Ile Glu Ile Pro Ile Thr Lys Gly Phe Pro Val
                    590                      595                      600
Leu Leu His Tyr Gln Thr Val Ser Glu Pro Ala Val Ile Lys Arg
                    605                      610                      615
Leu Ile Ser Val Leu Asn Lys Ser Thr Gly Glu Val Thr Lys Lys
                    620                      625                      630
Lys Pro Lys Phe Leu Thr Lys Gly Gln Asn Ala Leu Val Glu Leu
                    635                      640                      645
```

```
Gln Thr Gln Arg Pro Ile Ala Leu Glu Leu Tyr Lys Asp Phe Lys
                650                 655                 660
Glu Leu Gly Arg Phe Met Leu Arg Tyr Gly Gly Ser Thr Ile Ala
                665                 670                 675
Ala Gly Val Val Thr Glu Ile Lys Glu
                680
```

<210> 36
<211> 366
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2825460CD1

<400> 36

```
Met Ala Ala Ala Ala Ala Arg Trp Asn His Val Trp Val Gly Thr
  1               5                  10                  15
Glu Thr Gly Ile Leu Lys Gly Val Asn Leu Gln Arg Lys Gln Ala
                 20                  25                  30
Ala Asn Phe Thr Ala Gly Gly Gln Pro Arg Arg Glu Glu Ala Val
                 35                  40                  45
Ser Ala Leu Cys Trp Gly Thr Gly Gly Glu Thr Gln Met Leu Val
                 50                  55                  60
Gly Cys Ala Asp Arg Thr Val Lys His Phe Ser Thr Glu Asp Gly
                 65                  70                  75
Ile Phe Gln Gly Gln Arg His Cys Pro Gly Gly Glu Gly Met Phe
                 80                  85                  90
Arg Gly Leu Ala Gln Ala Asp Gly Thr Leu Ile Thr Cys Val Asp
                 95                 100                 105
Ser Gly Ile Leu Arg Val Trp His Asp Lys Asp Lys Asp Thr Ser
                110                 115                 120
Ser Asp Pro Leu Leu Glu Leu Arg Val Gly Pro Gly Val Cys Arg
                125                 130                 135
Met Arg Gln Asp Pro Ala His Pro His Val Val Ala Thr Gly Gly
                140                 145                 150
Lys Glu Asn Ala Leu Lys Ile Trp Asp Leu Gln Gly Ser Glu Glu
                155                 160                 165
Pro Val Phe Arg Ala Lys Asn Val Arg Asn Asp Trp Leu Asp Leu
                170                 175                 180
Arg Val Pro Ile Trp Asp Gln Asp Ile Gln Phe Leu Pro Gly Ser
                185                 190                 195
Gln Lys Leu Val Thr Cys Thr Gly Tyr His Gln Val Arg Val Tyr
                200                 205                 210
Asp Pro Ala Ser Pro Gln Arg Arg Pro Val Leu Glu Thr Thr Tyr
                215                 220                 225
Gly Glu Tyr Pro Leu Thr Ala Met Thr Leu Thr Pro Gly Gly Asn
                230                 235                 240
Ser Val Ile Val Gly Asn Thr His Gly Gln Leu Ala Glu Ile Asp
                245                 250                 255
Leu Arg Gln Gly Arg Leu Leu Gly Cys Leu Lys Gly Leu Ala Gly
                260                 265                 270
Ser Val Arg Gly Leu Gln Cys His Pro Ser Lys Pro Leu Leu Ala
                275                 280                 285
Ser Cys Gly Leu Asp Arg Val Leu Arg Ile His Arg Ile Gln Asn
                290                 295                 300
Pro Arg Gly Leu Glu His Lys Asp Glu Pro Gln Glu Pro Gln Glu
                305                 310                 315
Pro Asn Lys Val Pro Leu Glu Asp Thr Glu Thr Asp Glu Leu Trp
                320                 325                 330
Ala Ser Leu Glu Ala Ala Ala Lys Arg Lys Leu Ser Gly Leu Glu
                335                 340                 345
Gln Pro Gln Gly Ala Leu Gln Thr Arg Arg Arg Lys Lys Lys Arg
```

```
                350                 355                 360
Pro Gly Ser Thr Ser Pro
                365
<210> 37
<211> 339
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2871116CD1

<400> 37
Met Ala Thr Glu Ile Gly Ser Pro Pro Arg Phe Phe His Met Pro
  1               5                  10                  15
Arg Phe Gln His Gln Ala Pro Arg Gln Leu Phe Tyr Lys Arg Pro
                 20                  25                  30
Asp Phe Ala Gln Gln Gln Ala Met Gln Gln Leu Thr Phe Asp Gly
                 35                  40                  45
Lys Arg Met Arg Lys Ala Val Asn Arg Lys Thr Ile Asp Tyr Asn
                 50                  55                  60
Pro Ser Val Ile Lys Tyr Leu Glu Asn Arg Ile Trp Gln Arg Asp
                 65                  70                  75
Gln Arg Asp Met Arg Ala Ile Gln Pro Asp Ala Gly Tyr Tyr Asn
                 80                  85                  90
Asp Leu Val Pro Pro Ile Gly Met Leu Asn Asn Pro Met Asn Ala
                 95                 100                 105
Val Thr Thr Lys Phe Val Arg Thr Ser Thr Asn Lys Val Lys Cys
                110                 115                 120
Pro Val Phe Val Val Arg Leu Gln Glu Glu Phe Glu Ser Leu Ser
                125                 130                 135
Val Leu Lys Ser Trp Thr Pro Glu Gly Arg Arg Leu Val Thr Gly
                140                 145                 150
Ala Ser Ser Gly Glu Phe Thr Leu Trp Asn Gly Leu Thr Phe Asn
                155                 160                 165
Phe Glu Thr Ile Leu Gln Ala His Asp Ser Pro Val Arg Ala Met
                170                 175                 180
Thr Trp Ser His Asn Asp Met Trp Met Leu Thr Ala Asp His Gly
                185                 190                 195
Gly Tyr Val Lys Tyr Trp Gln Ser Asn Met Asn Asn Val Lys Met
                200                 205                 210
Phe Gln Ala His Lys Glu Ala Ile Arg Glu Ala Arg Phe Ile His
                215                 220                 225
Asn Ile Pro Phe Ser Val Val Pro Ile Val Met Val Lys Leu Phe
                230                 235                 240
Ser Lys Cys Ile Leu Gly Ala Glu Met His Gly Leu Cys Gln Phe
                245                 250                 255
Leu Gly Asn Phe Leu His Pro Ile Asn Thr Ile Phe Phe Phe Val
                260                 265                 270
Phe Thr His Ser Pro Phe Cys Trp His Leu Ser Glu Val Val Leu
                275                 280                 285
Ser Arg Tyr Gln Pro Leu Gln Tyr Val Arg Asp Val Leu Ser Ala
                290                 295                 300
Ala Phe Cys Thr Gly Phe Leu Phe Ser Phe Met Ile Asn Asn Val
                305                 310                 315
Tyr Thr Leu Phe Leu Phe Ile Ile Tyr Cys Val Arg Gln Glu Tyr
                320                 325                 330
Phe Ile Pro Asn Lys Glu Phe Ser Leu
                335
<210> 38
<211> 213
<212> PRT
<213> Homo sapiens
```

```
<220>
<221> misc_feature
<223> Incyte ID No: 2942212CD1

<400> 38
Met Glu Ala Ile Trp Leu Tyr Gln Phe Arg Leu Ile Val Ile Gly
 1               5                  10                  15
Asp Ser Thr Val Gly Lys Ser Cys Leu Ile Arg Arg Phe Thr Glu
                20                  25                  30
Gly Arg Phe Ala Gln Val Ser Asp Pro Thr Val Gly Val Asp Phe
                35                  40                  45
Phe Ser Arg Leu Val Glu Ile Glu Pro Gly Lys Arg Ile Lys Leu
                50                  55                  60
Gln Ile Trp Asp Thr Ala Gly Gln Glu Arg Phe Arg Ser Ile Thr
                65                  70                  75
Arg Ala Tyr Tyr Arg Asn Ser Val Gly Gly Leu Leu Leu Phe Ala
                80                  85                  90
Ile Thr Asn Arg Arg Ser Phe Gln Asn Val His Glu Trp Leu Glu
                95                 100                 105
Glu Thr Lys Val His Val Gln Pro Tyr Gln Ile Val Phe Val Leu
               110                 115                 120
Val Gly His Lys Cys Asp Leu Asp Thr Gln Arg Gln Val Thr Arg
               125                 130                 135
His Glu Ala Glu Lys Leu Ala Ala Ala Tyr Gly Met Lys Tyr Ile
               140                 145                 150
Glu Thr Ser Ala Arg Asp Ala Ile Asn Val Glu Lys Ala Phe Thr
               155                 160                 165
Asp Leu Thr Arg Asp Ile Tyr Glu Leu Val Lys Arg Gly Glu Ile
               170                 175                 180
Thr Ile Gln Glu Gly Trp Glu Gly Val Lys Ser Gly Phe Val Pro
               185                 190                 195
Asn Val Val His Ser Ser Glu Glu Val Val Lys Ser Glu Arg Arg
               200                 205                 210
Cys Leu Cys

<210> 39
<211> 393
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3685151CD1

<400> 39
Met Glu Leu Val Ala Gly Cys Tyr Glu Gln Val Leu Phe Gly Phe
 1               5                  10                  15
Ala Val His Pro Glu Pro Glu Ala Cys Gly Asp His Glu Gln Gln
                20                  25                  30
Trp Thr Leu Val Ala Asp Phe Thr His His Ala His Thr Ala Ser
                35                  40                  45
Leu Ser Ala Val Ala Val Asn Ser Arg Phe Val Val Thr Gly Ser
                50                  55                  60
Lys Asp Glu Thr Ile His Ile Tyr Asp Met Lys Lys Lys Ile Glu
                65                  70                  75
His Gly Ala Leu Val His His Ser Gly Thr Ile Thr Cys Leu Thr
                80                  85                  90
Phe Tyr Gly Asn Arg His Leu Ile Ser Gly Ala Glu Asp Gly Leu
                95                 100                 105
Ile Cys Ile Trp Asp Ala Lys Lys Trp Glu Ser Leu Thr Ser Ile
               110                 115                 120
Lys Ala His Lys Gly Gln Val Thr Phe Leu Ser Ile His Pro Ser
               125                 130                 135
```

Gly Lys Leu Ala Leu Ser Val Gly Thr Asp Lys Thr Leu Arg Thr
140 145 150
Trp Asn Leu Val Glu Gly Arg Ser Ala Phe Ile Lys Asn Ile Lys
155 160 165
Gln Asn Ala His Ile Val Glu Trp Ser Pro Arg Gly Glu Gln Tyr
170 175 180
Val Val Ile Ile Gln Asn Lys Ile Asp Ile Tyr Gln Leu Asp Thr
185 190 195
Ala Ser Ile Ser Gly Thr Ile Thr Asn Glu Lys Arg Ile Ser Ser
200 205 210
Val Lys Phe Leu Ser Glu Ser Val Leu Ala Val Ala Gly Asp Glu
215 220 225
Glu Val Ile Arg Phe Phe Asp Cys Asp Ser Leu Val Cys Leu Cys
230 235 240
Glu Phe Lys Ala His Glu Asn Arg Val Lys Asp Met Phe Ser Phe
245 250 255
Glu Ile Pro Glu His His Val Ile Val Ser Ala Ser Ser Asp Gly
260 265 270
Phe Ile Lys Met Trp Lys Leu Lys Gln Asp Lys Lys Val Pro Pro
275 280 285
Ser Leu Leu Cys Glu Ile Asn Thr Asn Ala Arg Leu Thr Cys Leu
290 295 300
Gly Val Trp Leu Asp Lys Val Ala Asp Met Lys Glu Ser Leu Pro
305 310 315
Pro Ala Ala Glu Pro Ser Pro Val Ser Lys Glu Gln Ser Lys Ile
320 325 330
Gly Lys Lys Glu Pro Gly Asp Thr Val His Lys Glu Glu Lys Arg
335 340 345
Ser Lys Pro Asn Thr Lys Lys Arg Gly Leu Thr Gly Asp Ser Lys
350 355 360
Lys Ala Thr Lys Glu Ser Gly Leu Ile Ser Thr Lys Lys Arg Lys
365 370 375
Met Val Glu Met Leu Glu Lys Lys Arg Lys Lys Lys Lys Ile Lys
380 385 390
Thr Met Gln

<210> 40
<211> 399
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4881515CD1

<400> 40
Met Ser Leu Gln Tyr Gly Ala Glu Glu Thr Pro Leu Ala Gly Ser
1 5 10 15
Tyr Gly Ala Ala Asp Ser Phe Pro Lys Asp Phe Gly Tyr Gly Val
20 25 30
Glu Glu Glu Glu Glu Glu Ala Ala Ala Ala Gly Gly Gly Val Gly
35 40 45
Ala Gly Ala Gly Gly Gly Cys Gly Pro Gly Gly Ala Asp Ser Ser
50 55 60
Lys Pro Arg Ile Leu Leu Met Gly Leu Arg Arg Ser Gly Lys Ser
65 70 75
Ser Ile Gln Lys Val Val Phe His Lys Met Ser Pro Asn Glu Thr
80 85 90
Leu Phe Leu Glu Ser Thr Asn Lys Ile Tyr Lys Asp Asp Ile Ser
95 100 105
Asn Ser Ser Phe Val Asn Phe Gln Ile Trp Asp Phe Pro Gly Gln
110 115 120
Met Asp Phe Phe Asp Pro Thr Phe Asp Tyr Glu Met Ile Phe Arg

```
                125                           130                           135
Gly Thr Gly Ala Leu Ile Tyr Val Ile Asp Ala Gln Asp Asp Tyr
                140                           145                           150
Met Glu Ala Leu Thr Arg Leu His Ile Thr Val Ser Lys Ala Tyr
                155                           160                           165
Lys Val Asn Pro Asp Met Asn Phe Glu Val Phe Ile His Lys Val
                170                           175                           180
Asp Gly Leu Ser Asp Asp His Lys Ile Glu Thr Gln Arg Asp Ile
                185                           190                           195
His Gln Arg Ala Asn Asp Asp Leu Ala Asp Ala Gly Leu Glu Lys
                200                           205                           210
Leu His Leu Ser Phe Tyr Leu Thr Ser Ile Tyr Asp His Ser Ile
                215                           220                           225
Phe Glu Ala Phe Ser Lys Val Val Gln Lys Leu Ile Pro Gln Leu
                230                           235                           240
Pro Thr Leu Glu Asn Leu Leu Asn Ile Phe Ile Ser Asn Ser Gly
                245                           250                           255
Ile Glu Lys Ala Phe Leu Phe Asp Val Val Ser Lys Ile Tyr Ile
                260                           265                           270
Ala Thr Asp Ser Ser Pro Val Asp Met Gln Ser Tyr Glu Leu Cys
                275                           280                           285
Cys Asp Met Ile Asp Val Val Ile Asp Val Ser Cys Ile Tyr Gly
                290                           295                           300
Leu Lys Glu Asp Gly Ser Gly Ser Ala Tyr Asp Lys Glu Ser Met
                305                           310                           315
Ala Ile Ile Lys Leu Asn Asn Thr Thr Val Leu Tyr Leu Lys Glu
                320                           325                           330
Val Thr Lys Phe Leu Ala Leu Val Cys Ile Leu Arg Glu Glu Ser
                335                           340                           345
Phe Glu Arg Lys Gly Leu Ile Asp Tyr Asn Phe His Cys Phe Arg
                350                           355                           360
Lys Ala Ile His Glu Val Phe Glu Val Gly Val Thr Ser His Arg
                365                           370                           375
Ser Cys Gly His Gln Thr Ser Ala Ser Ser Leu Lys Ala Leu Thr
                380                           385                           390
His Asn Gly Thr Pro Arg Asn Ala Ile
                395
```

<210> 41
<211> 412
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5324681CD1

<400> 41

```
Met Ala Gly Ser Val Gly Leu Ala Leu Cys Gly Gln Thr Leu Val
 1               5                  10                  15
Val Arg Gly Gly Ser Arg Phe Leu Ala Thr Ser Ile Ala Ser Ser
                20                  25                  30
Asp Asp Asp Ser Leu Phe Ile Tyr Asp Cys Ser Ala Ala Glu Lys
                35                  40                  45
Lys Ser Gln Glu Asn Lys Gly Glu Asp Ala Pro Leu Asp Gln Gly
                50                  55                  60
Ser Gly Ala Ile Leu Ala Ser Thr Phe Ser Lys Ser Gly Ser Tyr
                65                  70                  75
Phe Ala Leu Thr Asp Asp Ser Lys Arg Leu Ile Leu Phe Arg Thr
                80                  85                  90
Lys Pro Trp Gln Cys Leu Ser Val Arg Thr Val Ala Arg Arg Cys
                95                 100                 105
Thr Ala Leu Thr Phe Ile Ala Ser Glu Glu Lys Val Leu Val Ala
               110                 115                 120
```

```
Asp Lys Ser Gly Asp Val Tyr Ser Phe Ser Val Leu Glu Pro His
                125                 130                 135
Gly Cys Gly Arg Leu Glu Leu Gly His Leu Ser Met Leu Leu Asp
                140                 145                 150
Val Ala Val Ser Pro Asp Asp Arg Phe Ile Leu Thr Ala Asp Arg
                155                 160                 165
Asp Glu Lys Ile Arg Val Ser Trp Ala Ala Ala Pro His Ser Ile
                170                 175                 180
Glu Ser Phe Cys Leu Gly His Thr Glu Phe Val Ser Arg Ile Ser
                185                 190                 195
Val Val Pro Thr Gln Pro Gly Leu Leu Leu Ser Ser Ser Gly Asp
                200                 205                 210
Gly Thr Leu Arg Leu Trp Glu Tyr Arg Ser Gly Arg Gln Leu His
                215                 220                 225
Cys Cys His Leu Ala Ser Leu Gln Glu Leu Val Asp Pro Gln Ala
                230                 235                 240
Pro Gln Lys Phe Ala Ala Ser Arg Ile Ala Phe Trp Cys Gln Glu
                245                 250                 255
Asn Cys Val Ala Leu Leu Cys Asp Gly Thr Pro Val Val Tyr Ile
                260                 265                 270
Phe Gln Leu Asp Ala Arg Arg Gln Gln Leu Val Tyr Arg Gln Gln
                275                 280                 285
Leu Ala Phe Gln His Gln Val Trp Asp Val Ala Phe Glu Glu Thr
                290                 295                 300
Gln Gly Leu Trp Val Leu Gln Asp Cys Gln Glu Ala Pro Leu Val
                305                 310                 315
Leu Tyr Arg Pro Val Gly Asp Gln Trp Gln Ser Val Pro Glu Ser
                320                 325                 330
Thr Val Leu Lys Lys Val Ser Gly Val Leu Arg Gly Asn Trp Ala
                335                 340                 345
Met Leu Glu Gly Ser Ala Gly Ala Asp Ala Ser Phe Ser Ser Leu
                350                 355                 360
Tyr Lys Ala Thr Phe Asp Asn Val Thr Ser Tyr Leu Lys Lys Lys
                365                 370                 375
Glu Glu Arg Leu Gln Gln Gln Leu Glu Lys Lys Gln Arg Arg Arg
                380                 385                 390
Ser Pro Pro Pro Gly Pro Asp Gly His Ala Lys Lys Met Arg Pro
                395                 400                 405
Gly Glu Ala Thr Leu Ser Cys
                410
```

<210> 42
<211> 163
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5387651CD1

<400> 42

```
Met Asp Ala Leu Glu Gly Glu Ser Phe Ala Leu Ser Phe Ser Ser
  1               5                  10                  15
Ala Ser Asp Ala Glu Phe Asp Ala Val Val Gly Tyr Leu Glu Asp
                 20                  25                  30
Ile Ile Met Asp Asp Glu Phe Gln Leu Leu Gln Arg Asn Phe Met
                 35                  40                  45
Asp Lys Tyr Tyr Leu Glu Phe Glu Asp Thr Glu Glu Asn Lys Leu
                 50                  55                  60
Ile Tyr Thr Pro Ile Phe Asn Glu Tyr Ile Ser Leu Val Glu Lys
                 65                  70                  75
Tyr Ile Glu Glu Gln Leu Leu Gln Arg Ile Pro Glu Phe Asn Met
                 80                  85                  90
Ala Ala Phe Thr Thr Thr Leu Gln His His Lys Asp Glu Val Ala
```

```
                    95                  100                 105
Gly Asp Ile Phe Asp Met Leu Leu Thr Phe Thr Asp Phe Leu Ala
                110                 115                 120
Phe Lys Glu Met Phe Leu Asp Tyr Arg Ala Glu Lys Glu Gly Arg
                125                 130                 135
Gly Leu Asp Leu Ser Ser Gly Leu Val Val Thr Ser Leu Cys Lys
                140                 145                 150
Ser Ser Ser Leu Pro Ala Ser Gln Asn Asn Leu Arg His
                155                 160
```

<210> 43
<211> 514
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5595679CD1

<400> 43

```
Met Gln Glu Ser Gly Cys Arg Leu Glu His Pro Ser Ala Thr Lys
1               5                   10                  15
Phe Arg Asn His Val Met Glu Gly Asp Trp Asp Lys Ala Glu Asn
                20                  25                  30
Asp Leu Asn Glu Leu Lys Pro Leu Val His Ser Pro His Ala Ile
                35                  40                  45
Val Val Arg Gly Ala Leu Glu Ile Ser Gln Thr Leu Leu Gly Ile
                50                  55                  60
Ile Val Arg Met Lys Phe Leu Leu Leu Gln Gln Lys Tyr Leu Glu
                65                  70                  75
Tyr Leu Glu Asp Gly Lys Val Leu Glu Ala Leu Gln Val Leu Arg
                80                  85                  90
Cys Glu Leu Thr Pro Leu Lys Tyr Asn Thr Glu Arg Ile His Val
                95                  100                 105
Leu Ser Gly Tyr Leu Met Cys Ser His Ala Glu Asp Leu Arg Ala
                110                 115                 120
Lys Ala Glu Trp Glu Gly Lys Gly Thr Ala Ser Arg Ser Lys Leu
                125                 130                 135
Leu Asp Lys Leu Gln Thr Tyr Leu Pro Pro Ser Val Met Leu Pro
                140                 145                 150
Pro Arg Arg Leu Gln Thr Leu Leu Arg Gln Ala Val Glu Leu Gln
                155                 160                 165
Arg Asp Arg Cys Leu Tyr His Asn Thr Lys Leu Asp Asn Asn Leu
                170                 175                 180
Asp Ser Val Ser Leu Leu Ile Asp His Val Cys Ser Arg Arg Gln
                185                 190                 195
Phe Pro Cys Tyr Thr Gln Gln Ile Leu Thr Glu His Cys Asn Glu
                200                 205                 210
Val Trp Phe Cys Lys Phe Ser Asn Asp Gly Thr Lys Leu Ala Thr
                215                 220                 225
Gly Ser Lys Asp Thr Thr Val Ile Ile Trp Gln Val Asp Pro Asp
                230                 235                 240
Thr His Leu Leu Lys Leu Leu Lys Thr Leu Glu Gly His Ala Tyr
                245                 250                 255
Gly Val Ser Tyr Ile Ala Trp Ser Pro Asp Asp Asn Tyr Leu Val
                260                 265                 270
Ala Cys Gly Pro Asp Asp Cys Ser Glu Leu Trp Leu Trp Asn Val
                275                 280                 285
Gln Thr Gly Glu Leu Arg Thr Lys Met Ser Gln Ser His Glu Asp
                290                 295                 300
Ser Leu Thr Ser Val Ala Trp Asn Pro Asp Gly Lys Arg Phe Val
                305                 310                 315
Thr Gly Gly Gln Arg Gly Gln Phe Tyr Gln Cys Asp Leu Asp Gly
                320                 325                 330
```

Asn Leu Leu Asp Ser Trp Glu Gly Val Arg Val Gln Cys Leu Trp
335 340 345
Cys Leu Ser Asp Gly Lys Thr Val Leu Ala Ser Asp Thr His Gln
350 355 360
Arg Ile Arg Gly Tyr Asn Phe Glu Asp Leu Thr Asp Arg Asn Ile
365 370 375
Val Gln Glu Asp His Pro Ile Met Ser Phe Thr Ile Ser Lys Asn
380 385 390
Gly Arg Leu Ala Leu Leu Asn Val Ala Thr Gln Gly Val His Leu
395 400 405
Trp Asp Leu Gln Asp Arg Val Leu Val Arg Lys Tyr Gln Gly Val
410 415 420
Thr Gln Gly Phe Tyr Thr Ile His Ser Cys Phe Gly Gly His Asn
425 430 435
Glu Asp Phe Ile Ala Ser Gly Ser Glu Asp His Lys Val Tyr Ile
440 445 450
Trp His Lys Arg Ser Glu Leu Pro Ile Ala Glu Leu Thr Gly His
455 460 465
Thr Arg Thr Val Asn Cys Val Ser Trp Asn Pro Gln Ile Pro Ser
470 475 480
Met Met Ala Ser Ala Ser Asp Asp Gly Thr Val Arg Ile Trp Gly
485 490 495
Pro Ala Pro Phe Ile Asp His Gln Asn Ile Glu Glu Glu Cys Ser
500 505 510
Ser Met Asp Ser

<210> 44
<211> 67
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5782457CD1

<400> 44
Met Glu Glu Trp Asp Val Pro Gln Met Lys Lys Glu Val Glu Ser
1 5 10 15
Leu Lys Tyr Gln Leu Ala Phe Gln Arg Glu Met Ala Ser Lys Thr
20 25 30
Ile Pro Glu Leu Leu Lys Trp Ile Glu Asp Gly Ile Pro Lys Asp
35 40 45
Pro Phe Leu Asn Pro Asp Leu Met Lys Asn Asn Pro Trp Val Glu
50 55 60
Lys Gly Lys Cys Thr Ile Leu
65

<210> 45
<211> 315
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 760677CD1

<400> 45
Met Ala Phe Pro Glu Pro Lys Pro Arg Pro Pro Glu Leu Pro Gln
1 5 10 15
Lys Arg Leu Lys Thr Leu Asp Cys Gly Gln Gly Ala Val Arg Ala
20 25 30
Val Arg Phe Asn Val Asp Gly Asn Tyr Cys Leu Thr Cys Gly Ser
35 40 45
Asp Lys Thr Leu Lys Leu Trp Asn Pro Leu Arg Gly Thr Leu Leu

50 55 60
Arg Thr Tyr Ser Gly His Gly Tyr Glu Val Leu Asp Ala Ala Gly
65 70 75
Ser Phe Asp Asn Ser Ser Leu Cys Ser Gly Gly Gly Asp Lys Ala
80 85 90
Val Val Leu Trp Asn Val Ala Ser Gly Gln Val Val Arg Lys Phe
95 100 105
Arg Gly His Ala Gly Lys Val Asn Thr Val Gln Phe Ser Glu Glu
110 115 120
Ala Thr Val Ile Leu Ser Gly Ser Ile Asp Ser Ser Ile Arg Cys
125 130 135
Trp Asp Cys Arg Ser Arg Arg Pro Glu Pro Val Gln Thr Leu Asp
140 145 150
Glu Ala Arg Asp Gly Val Ser Ser Val Lys Val Ser Asp His Glu
155 160 165
Ile Leu Ala Gly Ser Val Asp Gly Arg Val Arg Arg Tyr Asp Leu
170 175 180
Arg Met Gly Gln Leu Phe Ser Asp Tyr Val Gly Ser Pro Ile Thr
185 190 195
Cys Thr Cys Phe Ser Arg Asp Gly Gln Cys Thr Leu Val Ser Ser
200 205 210
Leu Asp Ser Thr Leu Arg Leu Leu Asp Lys Asp Thr Gly Glu Leu
215 220 225
Leu Gly Glu Tyr Lys Gly His Lys Asn Gln Glu Tyr Lys Leu Asp
230 235 240
Cys Cys Leu Ser Glu Arg Asp Thr His Val Val Ser Cys Ser Glu
245 250 255
Asp Gly Lys Val Phe Phe Trp Asp Leu Val Glu Gly Ala Leu Ala
260 265 270
Leu Ala Leu Pro Val Gly Ser Gly Val Val Gln Ser Leu Asp Tyr
275 280 285
His Pro Thr Glu Pro Cys Leu Leu Thr Ala Met Gly Gly Ser Val
290 295 300
Gln Cys Trp Arg Glu Glu Ala Tyr Glu Ala Glu Asp Gly Ala Gly
305 310 315

<210> 46
<211> 504
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1348567CD1

<400> 46
Met Ser Leu Ile Cys Ser Ile Ser Asn Glu Val Pro Glu His Pro
1 5 10 15
Cys Val Ser Pro Val Ser Asn His Val Tyr Glu Arg Arg Leu Ile
20 25 30
Glu Lys Tyr Ile Ala Glu Asn Gly Thr Asp Pro Ile Asn Asn Gln
35 40 45
Pro Leu Ser Glu Glu Gln Leu Ile Asp Ile Lys Val Ala His Pro
50 55 60
Ile Arg Pro Lys Pro Pro Ser Ala Thr Ser Ile Pro Ala Ile Leu
65 70 75
Lys Ala Leu Gln Asp Glu Trp Asp Ala Val Met Pro His Ser Phe
80 85 90
Thr Leu Arg Gln Gln Leu Gln Thr Thr Arg Gln Glu Leu Ser His
95 100 105
Ala Leu Tyr Gln His Asp Ala Ala Cys Arg Val Ile Ala Arg Leu
110 115 120
Thr Lys Glu Val Thr Ala Ala Arg Glu Ala Leu Ala Thr Leu Lys

```
                    125                 130                 135
Pro Gln Ala Gly Leu Ile Val Pro Gln Ala Val Pro Ser Ser Gln
                    140                 145                 150
Pro Ser Val Val Gly Ala Gly Glu Pro Met Asp Leu Gly Glu Leu
                    155                 160                 165
Val Gly Met Thr Pro Glu Ile Ile Gln Lys Leu Gln Asp Lys Ala
                    170                 175                 180
Thr Val Leu Thr Thr Glu Arg Lys Lys Arg Gly Lys Thr Val Pro
                    185                 190                 195
Glu Glu Leu Val Lys Pro Glu Glu Leu Ser Lys Tyr Arg Gln Val
                    200                 205                 210
Ala Ser His Val Gly Leu His Ser Ala Ser Ile Pro Gly Ile Leu
                    215                 220                 225
Ala Leu Asp Leu Cys Pro Ser Asp Thr Asn Lys Ile Leu Thr Gly
                    230                 235                 240
Gly Ala Asp Lys Asn Val Val Val Phe Asp Lys Ser Ser Glu Gln
                    245                 250                 255
Ile Leu Ala Thr Leu Lys Gly His Thr Lys Lys Val Thr Ser Val
                    260                 265                 270
Val Phe His Pro Ser Gln Asp Leu Val Phe Ser Ala Ser Pro Asp
                    275                 280                 285
Ala Thr Ile Arg Ile Trp Ser Val Pro Asn Ala Ser Cys Val Gln
                    290                 295                 300
Val Val Arg Ala His Glu Ser Ala Val Thr Gly Leu Ser Leu His
                    305                 310                 315
Ala Thr Gly Asp Tyr Leu Leu Ser Ser Ser Asp Asp Gln Tyr Trp
                    320                 325                 330
Ala Phe Ser Asp Ile Gln Thr Gly Arg Val Leu Thr Lys Val Thr
                    335                 340                 345
Asp Glu Thr Ser Gly Cys Ser Leu Thr Cys Ala Gln Phe His Pro
                    350                 355                 360
Asp Gly Leu Ile Phe Gly Thr Gly Thr Met Asp Ser Gln Ile Lys
                    365                 370                 375
Ile Trp Asp Leu Lys Glu Arg Thr Asn Val Ala Asn Phe Pro Gly
                    380                 385                 390
His Ser Gly Pro Ile Thr Ser Ile Ala Phe Ser Glu Asn Gly Tyr
                    395                 400                 405
Tyr Leu Ala Thr Ala Ala Asp Asp Ser Ser Val Lys Leu Trp Asp
                    410                 415                 420
Leu Arg Lys Leu Lys Asn Phe Lys Thr Leu Gln Leu Asp Asn Asn
                    425                 430                 435
Phe Glu Val Lys Ser Leu Ile Phe Asp Gln Ser Gly Thr Tyr Leu
                    440                 445                 450
Ala Leu Gly Gly Thr Asp Val Gln Ile Tyr Ile Cys Lys Gln Trp
                    455                 460                 465
Thr Glu Ile Leu His Phe Thr Glu His Ser Gly Leu Thr Thr Gly
                    470                 475                 480
Val Ala Phe Gly His His Ala Lys Phe Ile Ala Ser Thr Gly Met
                    485                 490                 495
Asp Arg Ser Leu Lys Phe Tyr Ser Leu
                    500
```

<210> 47
<211> 522
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1751354CD1

<400> 47

```
Met Ala Phe Leu Asp Asn Pro Thr Ile Ile Leu Ala His Ile Arg
  1               5                  10                  15
```

Gln Ser His Val Thr Ser Asp Asp Thr Gly Met Cys Glu Met Val
20 25 30
Leu Ile Asp His Asp Val Asp Leu Glu Lys Ile His Pro Pro Ser
35 40 45
Met Pro Gly Asp Ser Gly Ser Glu Ile Gln Gly Ser Asn Gly Glu
50 55 60
Thr Gln Gly Tyr Val Tyr Ala Gln Ser Val Asp Ile Thr Ser Ser
65 70 75
Trp Asp Phe Gly Ile Arg Arg Arg Ser Asn Thr Ala Gln Arg Leu
80 85 90
Glu Arg Leu Arg Lys Glu Arg Gln Asn Gln Ile Lys Cys Lys Asn
95 100 105
Ile Gln Trp Lys Glu Arg Asn Ser Lys Gln Ser Ala Gln Glu Leu
110 115 120
Lys Ser Leu Phe Glu Lys Lys Ser Leu Lys Glu Lys Pro Pro Ile
125 130 135
Ser Gly Lys Gln Ser Ile Leu Ser Val Arg Leu Glu Gln Cys Pro
140 145 150
Leu Gln Leu Asn Asn Pro Phe Asn Glu Tyr Ser Lys Phe Asp Gly
155 160 165
Lys Gly His Val Gly Thr Thr Ala Thr Lys Lys Ile Asp Val Tyr
170 175 180
Leu Pro Leu His Ser Ser Gln Asp Arg Leu Leu Pro Met Thr Val
185 190 195
Val Thr Met Ala Ser Ala Arg Val Gln Asp Leu Ile Gly Leu Ile
200 205 210
Cys Trp Gln Tyr Thr Ser Glu Gly Arg Glu Pro Lys Leu Asn Asp
215 220 225
Asn Val Ser Ala Tyr Cys Leu His Ile Ala Glu Asp Asp Gly Glu
230 235 240
Val Asp Thr Asp Phe Pro Pro Leu Asp Ser Asn Glu Pro Ile His
245 250 255
Lys Phe Gly Phe Ser Thr Leu Ala Leu Val Glu Lys Tyr Ser Ser
260 265 270
Pro Gly Leu Thr Ser Lys Glu Ser Leu Phe Val Arg Ile Asn Ala
275 280 285
Ala His Gly Phe Ser Leu Ile Gln Val Asp Asn Thr Lys Val Thr
290 295 300
Met Lys Glu Ile Leu Leu Lys Ala Val Lys Arg Arg Lys Gly Ser
305 310 315
Gln Lys Val Ser Gly Pro Gln Tyr Arg Leu Glu Lys Gln Ser Glu
320 325 330
Pro Asn Val Ala Val Asp Leu Asp Ser Thr Leu Glu Ser Gln Ser
335 340 345
Ala Trp Glu Phe Cys Leu Val Arg Glu Asn Ser Ser Arg Ala Asp
350 355 360
Gly Val Phe Glu Glu Asp Ser Gln Ile Asp Ile Ala Thr Val Gln
365 370 375
Asp Met Leu Ser Ser His His Tyr Lys Ser Phe Lys Val Ser Met
380 385 390
Ile His Arg Leu Arg Phe Thr Thr Asp Val Gln Leu Gly Ile Ser
395 400 405
Gly Asp Lys Val Glu Ile Asp Pro Val Thr Asn Gln Lys Ala Ser
410 415 420
Thr Lys Phe Trp Ile Lys Gln Lys Pro Ile Ser Ile Asp Ser Asp
425 430 435
Leu Leu Cys Ala Cys Asp Leu Ala Glu Glu Lys Ser Pro Ser His
440 445 450
Ala Ile Phe Lys Leu Thr Tyr Leu Ser Asn His Asp Tyr Lys His
455 460 465
Leu Tyr Phe Glu Ser Asp Ala Ala Thr Val Asn Glu Ile Val Leu
470 475 480
Lys Val Asn Tyr Ile Leu Glu Ser Arg Ala Ser Thr Ala Arg Ala

485 490 495
Asp Tyr Phe Ala Gln Lys Gln Arg Lys Leu Asn Arg Arg Thr Ser
500 505 510
Phe Ser Phe Gln Lys Glu Lys Lys Ser Gly Gln Gln
515 520

<210> 48
<211> 316
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1976780CD1

<400> 48
Met Ala Ser Lys Asp Lys Ser Ser Lys Lys Asn Val Phe Glu Leu
1 5 10 15
Lys Thr Arg Gln Gly Thr Glu Leu Leu Ile Gln Ser Asp Asn Asp
20 25 30
Thr Val Ile Asn Asp Trp Phe Lys Val Leu Ser Ser Thr Ile Asn
35 40 45
Asn Gln Ala Val Glu Thr Asp Glu Gly Ile Glu Glu Glu Ile Pro
50 55 60
Asp Ser Pro Gly Ile Glu Lys His Asp Lys Glu Lys Glu Gln Lys
65 70 75
Asp Pro Lys Lys Leu Arg Ser Phe Lys Val Ser Ser Ile Asp Ser
80 85 90
Ser Glu Gln Lys Lys Thr Lys Lys Asn Leu Lys Lys Phe Leu Thr
95 100 105
Arg Arg Pro Thr Leu Gln Ala Val Arg Glu Lys Gly Tyr Ile Lys
110 115 120
Asp Gln Val Phe Gly Ser Asn Leu Ala Asn Leu Cys Gln Arg Glu
125 130 135
Asn Gly Thr Val Pro Lys Phe Val Lys Leu Cys Ile Glu His Val
140 145 150
Glu Glu His Gly Leu Asp Ile Asp Gly Ile Tyr Arg Val Ser Gly
155 160 165
Asn Leu Ala Val Ile Gln Lys Leu Arg Phe Ala Val Asn His Asp
170 175 180
Glu Lys Leu Asp Leu Asn Asp Ser Lys Trp Glu Asp Ile His Val
185 190 195
Ile Thr Gly Ala Leu Lys Met Phe Phe Arg Glu Leu Pro Glu Pro
200 205 210
Leu Phe Thr Phe Asn His Phe Asn Asp Phe Val Asn Ala Ile Lys
215 220 225
Gln Glu Pro Arg Gln Arg Val Ala Ala Val Lys Asp Leu Ile Arg
230 235 240
Gln Leu Pro Lys Pro Asn Gln Asp Thr Met Gln Ile Leu Phe Arg
245 250 255
His Leu Arg Arg Val Ile Glu Asn Gly Glu Lys Asn Arg Met Thr
260 265 270
Tyr Gln Ser Ile Ala Ile Val Phe Gly Pro Thr Leu Leu Lys Pro
275 280 285
Glu Lys Glu Thr Gly Asn Ile Ala Val His Thr Val Tyr Gln Asn
290 295 300
Gln Ile Val Glu Leu Ile Leu Leu Glu Leu Ser Ser Ile Phe Gly
305 310 315
Arg

<210> 49
<211> 387
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2048234CD1

<400> 49
Met Val His Cys Ser Cys Val Leu Phe Arg Lys Tyr Gly Asn Phe
1 5 10 15
Ile Asp Lys Leu Arg Leu Phe Thr Arg Gly Gly Ser Gly Gly Met
20 25 30
Gly Tyr Pro Arg Leu Gly Gly Glu Gly Gly Lys Gly Gly Asp Val
35 40 45
Trp Val Val Ala Gln Asn Arg Met Thr Leu Lys Gln Leu Lys Asp
50 55 60
Arg Tyr Pro Arg Lys Arg Phe Val Ala Gly Val Gly Ala Asn Ser
65 70 75
Lys Ile Ser Ala Leu Lys Gly Ser Lys Gly Lys Asp Trp Glu Ile
80 85 90
Pro Val Pro Val Gly Ile Ser Val Thr Asp Glu Asn Gly Lys Ile
95 100 105
Ile Gly Glu Leu Ser Lys Glu Asn Asp Arg Ile Leu Val Ala Gln
110 115 120
Gly Gly Leu Gly Gly Lys Leu Leu Thr Asn Phe Leu Pro Leu Lys
125 130 135
Gly Gln Lys Arg Ile Ile His Leu Asp Leu Lys Leu Ile Ala Asp
140 145 150
Val Gly Leu Val Gly Phe Pro Asn Ala Gly Lys Ser Ser Leu Leu
155 160 165
Ser Cys Val Ser His Ala Lys Pro Ala Ile Ala Asp Tyr Ala Phe
170 175 180
Thr Thr Leu Lys Leu Lys Leu Gly Lys Ile Met Tyr Ser Asp Phe
185 190 195
Lys Gln Ile Ser Val Ala Asp Leu Pro Gly Leu Ile Glu Gly Ala
200 205 210
His Met Asn Lys Gly Met Gly His Lys Phe Leu Lys His Ile Glu
215 220 225
Arg Thr Arg Gln Leu Leu Phe Val Val Asp Ile Ser Gly Phe Gln
230 235 240
Leu Ser Ser His Thr Gln Tyr Arg Thr Ala Phe Glu Thr Ile Ile
245 250 255
Leu Leu Thr Lys Glu Leu Glu Leu Tyr Lys Glu Glu Leu Gln Thr
260 265 270
Lys Pro Ala Leu Leu Ala Val Asn Lys Met Asp Leu Pro Asp Ala
275 280 285
Gln Asp Lys Phe His Glu Leu Met Ser Gln Leu Gln Asn Pro Lys
290 295 300
Asp Phe Leu His Leu Phe Glu Lys Asn Met Ile Pro Glu Arg Thr
305 310 315
Val Glu Phe Gln His Ile Ile Pro Ile Ser Ala Val Thr Gly Glu
320 325 330
Gly Ile Glu Glu Leu Lys Asn Cys Ile Arg Lys Ser Leu Asp Glu
335 340 345
Gln Ala Asn Gln Glu Asn Asp Ala Leu His Lys Lys Gln Leu Leu
350 355 360
Asn Leu Trp Ile Ser Asp Thr Met Ser Ser Thr Glu Pro Pro Ser
365 370 375
Lys His Ala Val Thr Thr Ser Lys Met Asp Ile Ile
380 385

<210> 50
<211> 334
<212> PRT
<213> Homo sapiens

<220>

<221> misc_feature
<223> Incyte ID No: 2111754CD1

<400> 50
Met Pro Ser Gly Pro Arg Ala Ala Leu Arg Trp Ala Ser Pro Ser
1 5 10 15
Gln Leu Val Ser Tyr His Val Leu Arg Asn Gly Ile Tyr Ala Cys
20 25 30
Tyr Pro His Ser Leu Arg Pro Arg Thr Pro Leu Leu Cys Ala Ser
35 40 45
Arg Asn Ile Lys Pro Arg Arg Ser Glu Leu Leu Gly Cys Pro Val
50 55 60
Gly Cys Arg Gly Ser Leu Ser Glu Gln Arg Ile Cys Leu Leu Gly
65 70 75
Cys Leu Val Arg Ala Ser Glu Lys Gly Val Ser Cys Cys Gln Leu
80 85 90
Ser Val Gly Glu Leu Val His Val Ser Pro Leu Arg Ile Pro Thr
95 100 105
Met Gly Asn Ala Ser Phe Gly Ser Lys Glu Gln Lys Leu Leu Lys
110 115 120
Arg Leu Arg Leu Leu Pro Ala Leu Leu Ile Leu Arg Ala Phe Lys
125 130 135
Pro His Arg Lys Ile Arg Asp Tyr Arg Val Val Val Val Gly Thr
140 145 150
Ala Gly Val Gly Lys Ser Thr Leu Leu His Lys Trp Ala Ser Gly
155 160 165
Asn Phe Arg His Glu Tyr Leu Pro Thr Ile Glu Asn Thr Tyr Cys
170 175 180
Gln Leu Leu Gly Cys Ser His Gly Val Leu Ser Leu His Ile Thr
185 190 195
Asp Ser Lys Ser Gly Asp Gly Asn Arg Ala Leu Gln Arg His Val
200 205 210
Ile Ala Arg Gly His Ala Phe Val Leu Val Tyr Ser Val Thr Lys
215 220 225
Lys Glu Thr Leu Glu Glu Leu Lys Ala Phe Tyr Glu Leu Ile Cys
230 235 240
Lys Ile Lys Gly Asn Asn Leu His Lys Phe Pro Ile Val Leu Val
245 250 255
Gly Asn Lys Ser Asp Asp Thr His Arg Glu Val Ala Leu Asn Asp
260 265 270
Gly Ala Thr Cys Ala Met Glu Trp Asn Cys Ala Phe Met Glu Ile
275 280 285
Ser Ala Lys Thr Asp Val Asn Val Gln Glu Leu Phe His Met Leu
290 295 300
Leu Asn Tyr Lys Lys Lys Pro Thr Thr Gly Leu Gln Glu Pro Glu
305 310 315
Lys Lys Ser Gln Met Pro Asn Thr Thr Glu Lys Leu Leu Asp Lys
320 325 330
Cys Ile Ile Met

<210> 51
<211> 551
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2123286CD1

<400> 51
Met Glu Glu Glu Leu Pro Leu Phe Ser Gly Asp Ser Gly Lys Pro
1 5 10 15
Val Gln Ala Thr Leu Ser Ser Leu Lys Met Leu Asp Val Gly Lys

20 25 30
Trp Pro Ile Phe Ser Leu Cys Ser Glu Glu Glu Leu Gln Leu Ile
35 40 45
Arg Gln Ala Cys Val Phe Gly Ser Ala Gly Asn Glu Val Leu Tyr
50 55 60
Thr Thr Val Asn Asp Glu Ile Phe Val Leu Gly Thr Asn Cys Cys
65 70 75
Gly Cys Leu Gly Leu Gly Asp Val Gln Ser Thr Ile Glu Pro Arg
80 85 90
Arg Leu Asp Ser Leu Asn Gly Lys Lys Ile Ala Cys Leu Ser Tyr
95 100 105
Gly Ser Gly Pro His Ile Val Leu Ala Thr Thr Glu Gly Glu Val
110 115 120
Phe Thr Trp Gly His Asn Ala Tyr Ser Gln Leu Gly Asn Gly Thr
125 130 135
Thr Asn His Gly Leu Val Pro Cys His Ile Ser Thr Asn Leu Ser
140 145 150
Asn Lys Gln Val Ile Glu Val Ala Cys Gly Ser Tyr His Ser Leu
155 160 165
Val Leu Thr Ser Asp Gly Glu Val Phe Ala Trp Gly Tyr Asn Asn
170 175 180
Ser Gly Gln Val Gly Ser Gly Ser Thr Val Asn Gln Pro Ile Pro
185 190 195
Arg Arg Val Thr Gly Cys Leu Gln Asn Lys Val Val Val Thr Ile
200 205 210
Ala Cys Gly Gln Met Cys Cys Met Ala Val Val Asp Thr Gly Glu
215 220 225
Val Tyr Val Trp Gly Tyr Asn Gly Asn Gly Gln Leu Gly Leu Gly
230 235 240
Asn Ser Gly Asn Gln Pro Thr Pro Cys Arg Val Ala Ala Leu Gln
245 250 255
Gly Ile Arg Val Gln Arg Val Ala Cys Gly Tyr Ala His Thr Leu
260 265 270
Val Leu Thr Asp Glu Gly Gln Val Tyr Ala Trp Gly Ala Asn Ser
275 280 285
Tyr Gly Gln Leu Gly Thr Gly Asn Lys Ser Asn Gln Ser Tyr Pro
290 295 300
Thr Pro Val Thr Val Glu Lys Asp Arg Ile Ile Glu Ile Ala Ala
305 310 315
Cys His Ser Thr His Thr Ser Ala Ala Lys Thr Gln Gly Gly His
320 325 330
Val Tyr Met Trp Gly Gln Cys Arg Gly Gln Ser Val Ile Leu Pro
335 340 345
His Leu Thr His Phe Ser Cys Thr Asp Asp Val Phe Ala Cys Phe
350 355 360
Ala Thr Pro Ala Val Thr Trp Arg Leu Leu Ser Val Glu Pro Asp
365 370 375
Asp His Leu Thr Val Ala Glu Ser Leu Lys Arg Glu Phe Asp Asn
380 385 390
Pro Asp Thr Ala Asp Leu Lys Phe Leu Val Asp Gly Lys Tyr Ile
395 400 405
Tyr Ala His Lys Val Leu Leu Lys Ile Arg Cys Glu His Phe Arg
410 415 420
Ser Ser Leu Glu Asp Asn Glu Asp Asp Ile Val Glu Met Ser Glu
425 430 435
Phe Ser Tyr Pro Val Tyr Arg Ala Phe Leu Glu Tyr Leu Tyr Thr
440 445 450
Asp Ser Ile Ser Leu Ser Pro Glu Glu Ala Val Gly Leu Leu Asp
455 460 465
Leu Ala Thr Phe Tyr Arg Glu Asn Arg Leu Lys Lys Leu Cys Gln
470 475 480
Gln Thr Ile Lys Gln Gly Ile Cys Glu Glu Asn Ala Ile Ala Leu
485 490 495

Leu Ser Ala Ala Val Lys Tyr Asp Ala Gln Asp Leu Glu Glu Phe
500 505 510
Cys Phe Arg Phe Cys Ile Asn His Leu Thr Val Val Thr Gln Thr
515 520 525
Ser Gly Phe Ala Glu Met Asp His Asp Leu Leu Lys Asn Phe Ile
530 535 540
Ser Lys Ala Ser Arg Val Gly Ala Phe Lys Asn
545 550

<210> 52
<211> 308
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2477507CD1

<400> 52
Met Ile His Asp Ala Gln Met Asp Tyr Tyr Gly Thr Arg Leu Ala
1 5 10 15
Thr Cys Ser Ser Asp Arg Ser Val Lys Ile Phe Asp Val Arg Asn
20 25 30
Gly Gly Gln Ile Leu Ile Ala Asp Leu Arg Gly His Glu Gly Pro
35 40 45
Val Trp Gln Val Ala Trp Ala His Pro Met Tyr Gly Asn Ile Leu
50 55 60
Ala Ser Cys Ser Tyr Asp Arg Lys Val Ile Ile Trp Arg Glu Glu
65 70 75
Asn Gly Thr Trp Glu Lys Ser His Glu His Ala Gly His Asp Ser
80 85 90
Ser Val Asn Ser Val Cys Trp Ala Pro His Asp Tyr Gly Leu Ile
95 100 105
Leu Ala Cys Gly Ser Ser Asp Gly Ala Ile Ser Leu Leu Thr Tyr
110 115 120
Thr Gly Glu Gly Gln Trp Glu Val Lys Lys Ile Asn Asn Ala His
125 130 135
Thr Ile Gly Cys Asn Ala Val Ser Trp Ala Pro Ala Val Val Pro
140 145 150
Gly Ser Leu Ile Asp His Pro Ser Gly Gln Lys Pro Asn Tyr Ile
155 160 165
Lys Arg Phe Ala Ser Gly Gly Cys Asp Asn Leu Ile Lys Leu Trp
170 175 180
Lys Glu Glu Glu Asp Gly Gln Trp Lys Glu Glu Gln Lys Leu Glu
185 190 195
Ala His Ser Asp Trp Val Arg Asp Val Ala Trp Ala Pro Ser Ile
200 205 210
Gly Leu Pro Thr Ser Thr Ile Ala Ser Cys Ser Gln Asp Gly Arg
215 220 225
Val Phe Ile Trp Thr Cys Asp Asp Ala Ser Ser Asn Thr Trp Ser
230 235 240
Pro Lys Leu Leu His Lys Phe Asn Asp Val Val Trp His Val Ser
245 250 255
Trp Ser Ile Thr Ala Asn Ile Leu Ala Val Ser Gly Gly Asp Asn
260 265 270
Lys Val Thr Leu Trp Lys Glu Ser Val Asp Gly Gln Trp Val Cys
275 280 285
Ile Ser Asp Val Asn Lys Gly Gln Gly Ser Val Ser Ala Ser Val
290 295 300
Thr Glu Gly Gln Gln Asn Glu Gln
305

<210> 53
<211> 949
<212> PRT

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2759119CD1

<400> 53
```
Met Asp Ala Leu Glu Asp Tyr Val Trp Pro Arg Ala Thr Ser Glu
 1               5                  10                  15
Leu Ile Leu Leu Pro Val Thr Gly Leu Glu Cys Val Gly Asp Arg
                20                  25                  30
Leu Leu Ala Gly Glu Gly Pro Asp Val Leu Val Tyr Ser Leu Asp
                35                  40                  45
Phe Gly Gly His Leu Arg Met Ile Lys Arg Val Gln Asn Leu Leu
                50                  55                  60
Gly His Tyr Leu Ile His Gly Phe Arg Val Arg Pro Glu Pro Asn
                65                  70                  75
Gly Asp Leu Asp Leu Glu Ala Met Val Ala Val Phe Gly Ser Lys
                80                  85                  90
Gly Leu Arg Val Val Lys Ile Ser Trp Gly Gln Gly His Phe Trp
                95                 100                 105
Glu Leu Trp Arg Ser Gly Leu Trp Asn Met Ser Asp Trp Ile Trp
               110                 115                 120
Asp Ala Arg Trp Leu Glu Gly Asn Ile Ala Leu Ala Leu Gly His
               125                 130                 135
Asn Ser Val Val Leu Tyr Asp Pro Val Val Gly Cys Ile Leu Gln
               140                 145                 150
Glu Val Pro Cys Thr Asp Arg Cys Thr Leu Ser Ser Ala Cys Leu
               155                 160                 165
Ile Gly Asp Ala Trp Lys Glu Leu Thr Ile Val Ala Gly Ala Val
               170                 175                 180
Ser Asn Gln Leu Leu Val Trp Tyr Pro Ala Thr Ala Leu Ala Asp
               185                 190                 195
Asn Lys Pro Val Ala Pro Asp Arg Arg Ile Ser Gly His Val Gly
               200                 205                 210
Ile Ile Phe Ser Met Ser Tyr Leu Glu Ser Lys Gly Leu Leu Ala
               215                 220                 225
Thr Ala Ser Glu Asp Arg Ser Val Arg Ile Trp Lys Val Gly Asp
               230                 235                 240
Leu Arg Val Pro Gly Gly Arg Val Gln Asn Ile Gly His Cys Phe
               245                 250                 255
Gly His Ser Ala Arg Val Trp Gln Val Lys Leu Leu Glu Asn Tyr
               260                 265                 270
Leu Ile Ser Ala Gly Glu Asp Cys Val Cys Leu Val Trp Ser His
               275                 280                 285
Glu Gly Glu Ile Leu Gln Ala Phe Arg Gly His Gln Gly Arg Gly
               290                 295                 300
Ile Arg Ala Ile Ala Ala His Glu Arg Gln Ala Trp Val Ile Thr
               305                 310                 315
Gly Gly Asp Asp Ser Gly Ile Arg Leu Trp His Leu Val Gly Arg
               320                 325                 330
Gly Tyr Arg Gly Leu Gly Val Ser Ala Leu Cys Phe Lys Ser Arg
               335                 340                 345
Ser Arg Pro Gly Thr Leu Lys Ala Val Thr Leu Ala Gly Ser Trp
               350                 355                 360
Arg Leu Leu Ala Val Thr Asp Thr Gly Ala Leu Tyr Leu Tyr Asp
               365                 370                 375
Val Glu Val Lys Cys Trp Glu Gln Leu Leu Glu Asp Lys His Phe
               380                 385                 390
Gln Ser Tyr Cys Leu Leu Glu Ala Ala Pro Gly Pro Glu Gly Phe
               395                 400                 405
Gly Leu Cys Ala Met Ala Asn Gly Glu Gly Arg Val Lys Val Val
               410                 415                 420
```

Pro Ile Asn Thr Pro Thr Ala Ala Val Asp Gln Thr Leu Phe Pro
425 430 435
Gly Lys Val His Ser Leu Ser Trp Ala Leu Arg Gly Tyr Glu Glu
440 445 450
Leu Leu Leu Leu Ala Ser Gly Pro Gly Gly Val Val Ala Cys Leu
455 460 465
Glu Ile Ser Ala Ala Pro Ser Gly Lys Ala Ile Phe Val Lys Glu
470 475 480
Arg Cys Arg Tyr Leu Leu Pro Pro Ser Lys Gln Arg Trp His Thr
485 490 495
Cys Ser Ala Phe Leu Pro Pro Gly Asp Phe Leu Val Cys Gly Asp
500 505 510
Arg Arg Gly Ser Val Leu Leu Phe Pro Ser Arg Pro Gly Leu Leu
515 520 525
Lys Asp Pro Gly Val Gly Gly Lys Ala Arg Ala Gly Ala Gly Ala
530 535 540
Pro Val Val Gly Ser Gly Ser Ser Gly Gly Gly Asn Ala Phe Thr
545 550 555
Gly Leu Gly Pro Val Ser Thr Leu Pro Ser Leu His Gly Lys Gln
560 565 570
Gly Val Thr Ser Val Thr Cys His Gly Gly Tyr Val Tyr Thr Ile
575 580 585
Gly Arg Asp Gly Ala Tyr Tyr Gln Leu Phe Val Arg Asp Gly Gln
590 595 600
Leu Gln Pro Val Leu Arg Gln Lys Ser Cys Arg Gly Met Asn Trp
605 610 615
Leu Ala Gly Leu Arg Ile Val Pro Asp Gly Ser Met Val Ile Leu
620 625 630
Gly Phe His Ala Asn Glu Phe Val Val Trp Asn Pro Arg Ser His
635 640 645
Glu Lys Leu His Ile Val Asn Cys Gly Gly Gly His Arg Ser Trp
650 655 660
Ala Phe Ser Asp Thr Glu Ala Ala Met Ala Phe Ala Tyr Leu Lys
665 670 675
Asp Gly Asp Val Met Leu Tyr Arg Ala Leu Gly Gly Cys Thr Arg
680 685 690
Pro His Val Ile Leu Arg Glu Gly Leu His Gly Arg Glu Ile Thr
695 700 705
Cys Val Lys Arg Val Gly Thr Ile Thr Leu Gly Pro Glu Tyr Gly
710 715 720
Val Pro Ser Phe Met Gln Pro Asp Asp Leu Glu Pro Gly Ser Glu
725 730 735
Gly Pro Asp Leu Thr Asp Ile Val Ile Thr Cys Ser Glu Asp Thr
740 745 750
Thr Val Cys Val Leu Ala Leu Pro Thr Thr Thr Gly Ser Ala His
755 760 765
Ala Leu Thr Ala Val Cys Asn His Ile Ser Ser Val Arg Ala Val
770 775 780
Ala Val Trp Gly Ile Gly Thr Pro Gly Gly Pro Gln Asp Pro Gln
785 790 795
Pro Gly Leu Thr Ala His Val Val Ser Ala Gly Gly Arg Ala Glu
800 805 810
Met His Cys Phe Ser Ile Met Val Thr Pro Asp Pro Ser Thr Pro
815 820 825
Ser Arg Leu Ala Cys His Val Met His Leu Ser Ser His Arg Leu
830 835 840
Asp Glu Tyr Trp Asp Arg Gln Arg Asn Arg His Arg Met Val Lys
845 850 855
Val Asp Pro Glu Thr Arg Tyr Met Ser Leu Ala Val Cys Glu Leu
860 865 870
Asp Gln Pro Gly Leu Gly Pro Leu Val Ala Ala Ala Cys Ser Asp
875 880 885
Gly Ala Val Ser Ser Phe Phe Cys Arg Ile Leu Gly Gly Phe Cys

```
                890                 895                 900
Ser Ser Leu Leu Lys Pro Ser Thr Ile Ser Asp Val Ser Ser Arg
                905                 910                 915
Ser Thr Pro Leu His Thr Arg His Pro Thr Arg Gly Gly Gly Ser
                920                 925                 930
Ser Cys Ala Ala Gln Leu Leu Met Ala Ala Trp Leu Ser Gly Ile
                935                 940                 945
Ser Pro Pro Cys

<210> 54
<211> 227
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2823818CD1

<400> 54
Met Arg His Glu Ala Pro Met Gln Met Ala Ser Ala Gln Asp Ala
 1               5                  10                  15
Arg Tyr Gly Gln Lys Asp Ser Ser Asp Gln Asn Phe Asp Tyr Met
                20                  25                  30
Phe Lys Leu Leu Ile Ile Gly Asn Ser Ser Val Gly Lys Thr Ser
                35                  40                  45
Phe Leu Phe Arg Tyr Ala Asp Asp Ser Phe Thr Ser Ala Phe Val
                50                  55                  60
Ser Thr Val Gly Ile Asp Phe Lys Val Lys Thr Val Phe Lys Asn
                65                  70                  75
Val Lys Arg Ile Lys Leu Gln Ile Trp Asp Thr Ala Gly Gln Glu
                80                  85                  90
Arg Tyr Arg Thr Ile Thr Thr Ala Tyr Tyr Arg Gly Ala Met Gly
                95                 100                 105
Phe Ile Leu Met Tyr Asp Ile Thr Asn Glu Glu Ser Phe Asn Ala
               110                 115                 120
Val Gln Asp Trp Ser Thr Gln Ile Lys Thr Tyr Ser Trp Asp Asn
               125                 130                 135
Ala Gln Val Ile Leu Val Gly Asn Lys Cys Asp Met Glu Asp Glu
               140                 145                 150
Arg Val Ile Ser Thr Glu Arg Gly Gln His Leu Gly Glu Gln Leu
               155                 160                 165
Gly Phe Glu Phe Phe Glu Thr Ser Ala Lys Asp Asn Ile Asn Val
               170                 175                 180
Lys Gln Thr Phe Glu Arg Leu Val Asp Ile Ile Cys Asp Lys Met
               185                 190                 195
Ser Glu Ser Leu Glu Thr Asp Pro Ala Ile Thr Ala Ala Lys Gln
               200                 205                 210
Asn Thr Arg Leu Lys Glu Thr Pro Pro Pro Pro Gln Pro Asn Cys
               215                 220                 225
Ala Cys

<210> 55
<211> 474
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2859730CD1

<400> 55
Met Arg Arg Val Val Arg Gln Ser Lys Phe Arg His Val Phe Gly
 1               5                  10                  15
```

```
Gln Ala Val Lys Asn Asp Gln Cys Tyr Asp Asp Ile Arg Val Ser
                20                  25                  30
Arg Val Thr Trp Asp Ser Ser Phe Cys Ala Val Asn Pro Arg Phe
                35                  40                  45
Val Ala Ile Ile Ile Glu Ala Ser Gly Gly Gly Ala Phe Leu Val
                50                  55                  60
Leu Pro Leu Arg Lys Thr Gly Arg Ile Asp Lys Ser Tyr Pro Thr
                65                  70                  75
Val Cys Gly His Thr Gly Pro Val Leu Asp Ile Asp Trp Cys Pro
                80                  85                  90
His Asn Asp Gln Val Ile Ala Ser Gly Ser Glu Asp Cys Thr Val
                95                  100                 105
Met Val Trp Gln Ile Pro Glu Asn Gly Leu Thr Leu Ser Leu Thr
                110                 115                 120
Glu Pro Val Val Ile Leu Glu Gly His Ser Lys Arg Val Gly Ile
                125                 130                 135
Val Ala Trp His Pro Thr Ala Arg Asn Val Leu Leu Ser Ala Gly
                140                 145                 150
Cys Asp Asn Ala Ile Ile Ile Trp Asn Val Gly Thr Gly Glu Ala
                155                 160                 165
Leu Ile Asn Leu Asp Asp Met His Ser Asp Met Ile Tyr Asn Val
                170                 175                 180
Ser Trp Asn Arg Asn Gly Ser Leu Ile Cys Thr Ala Ser Lys Asp
                185                 190                 195
Lys Lys Val Arg Val Ile Asp Pro Arg Lys Gln Glu Ile Val Ala
                200                 205                 210
Glu Lys Glu Lys Ala His Glu Gly Ala Arg Pro Met Arg Ala Ile
                215                 220                 225
Phe Leu Ala Asp Gly Asn Val Phe Thr Thr Gly Phe Ser Arg Met
                230                 235                 240
Ser Glu Arg Gln Leu Ala Leu Trp Asn Pro Lys Asn Met Gln Glu
                245                 250                 255
Pro Ile Ala Leu His Glu Met Asp Thr Ser Asn Gly Val Leu Leu
                260                 265                 270
Pro Phe Tyr Asp Pro Asp Thr Ser Ile Ile Tyr Leu Cys Gly Lys
                275                 280                 285
Gly Asp Ser Ser Ile Arg Tyr Phe Glu Ile Thr Asp Glu Ser Pro
                290                 295                 300
Tyr Val His Tyr Leu Asn Thr Phe Ser Ser Lys Glu Pro Gln Arg
                305                 310                 315
Gly Met Gly Tyr Met Pro Lys Arg Gly Leu Asp Val Asn Lys Cys
                320                 325                 330
Glu Ile Ala Arg Phe Phe Lys Leu His Glu Arg Lys Cys Glu Pro
                335                 340                 345
Ile Ile Met Thr Val Pro Arg Lys Ser Asp Leu Phe Gln Asp Asp
                350                 355                 360
Leu Tyr Pro Asp Thr Ala Gly Pro Glu Ala Ala Leu Glu Ala Glu
                365                 370                 375
Glu Trp Phe Glu Gly Lys Asn Ala Asp Pro Ile Leu Ile Ser Leu
                380                 385                 390
Lys His Gly Tyr Ile Pro Gly Lys Asn Arg Asp Leu Lys Val Val
                395                 400                 405
Lys Lys Asn Ile Leu Asp Ser Lys Pro Thr Ala Asn Lys Lys Cys
                410                 415                 420
Asp Leu Ile Ser Ile Pro Lys Lys Thr Thr Asp Thr Ala Ser Val
                425                 430                 435
Gln Asn Glu Ala Lys Leu Asp Glu Ile Leu Lys Glu Ile Lys Ser
                440                 445                 450
Ile Lys Asp Thr Ile Cys Asn Gln Asp Glu Arg Ile Ser Lys Leu
                455                 460                 465
Glu Gln Gln Met Ala Lys Ile Ala Ala
                470
```

<210> 56

<211> 547
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2861155CD1

<400> 56
```
Met Lys Thr Leu Glu Thr Gln Pro Leu Ala Pro Asp Cys Cys Pro
 1               5                  10                  15
Ser Asp Gln Asp Pro Ala Pro Ala His Pro Ser Pro His Ala Ser
                20                  25                  30
Pro Met Asn Lys Asn Ala Asp Ser Glu Leu Met Pro Pro Pro Pro
                35                  40                  45
Glu Arg Gly Asp Pro Pro Arg Leu Ser Pro Asp Pro Val Ala Gly
                50                  55                  60
Ser Ala Val Ser Gln Glu Leu Arg Glu Gly Asp Pro Val Ser Leu
                65                  70                  75
Ser Thr Pro Leu Glu Thr Glu Phe Gly Ser Pro Ser Glu Leu Ser
                80                  85                  90
Pro Arg Ile Glu Glu Gln Glu Leu Ser Glu Asn Thr Ser Leu Pro
                95                 100                 105
Ala Glu Glu Ala Asn Gly Ser Leu Ser Glu Glu Glu Ala Asn Gly
               110                 115                 120
Pro Glu Leu Gly Ser Gly Lys Ala Met Glu Asp Thr Ser Gly Glu
               125                 130                 135
Pro Ala Ala Glu Asp Glu Gly Asp Thr Ala Trp Asn Tyr Ser Phe
               140                 145                 150
Ser Gln Leu Pro Arg Phe Leu Ser Gly Ser Trp Ser Glu Phe Ser
               155                 160                 165
Thr Gln Pro Glu Asn Phe Leu Lys Gly Cys Lys Trp Ala Pro Asp
               170                 175                 180
Gly Ser Cys Ile Leu Thr Asn Ser Ala Asp Asn Ile Leu Arg Ile
               185                 190                 195
Tyr Asn Leu Pro Pro Glu Leu Tyr His Glu Gly Glu Gln Val Glu
               200                 205                 210
Tyr Ala Glu Met Val Pro Val Leu Arg Met Val Glu Gly Asp Thr
               215                 220                 225
Ile Tyr Asp Tyr Cys Trp Tyr Ser Leu Met Ser Ser Ala Gln Pro
               230                 235                 240
Asp Thr Ser Tyr Val Ala Ser Ser Ser Arg Glu Asn Pro Ile His
               245                 250                 255
Ile Trp Asp Ala Phe Thr Gly Glu Leu Arg Ala Ser Phe Arg Ala
               260                 265                 270
Tyr Asn His Leu Asp Glu Leu Thr Ala Ala His Ser Leu Cys Phe
               275                 280                 285
Ser Pro Asp Gly Ser Gln Leu Phe Cys Gly Phe Asn Arg Thr Val
               290                 295                 300
Arg Val Phe Ser Thr Ala Arg Pro Gly Arg Asp Cys Glu Val Arg
               305                 310                 315
Ala Thr Phe Ala Lys Lys Gln Gly Gln Ser Gly Ile Ile Ser Cys
               320                 325                 330
Ile Ala Phe Ser Pro Ala Gln Pro Leu Tyr Ala Cys Gly Ser Tyr
               335                 340                 345
Gly Arg Ser Leu Gly Leu Tyr Ala Trp Asp Asp Gly Ser Pro Leu
               350                 355                 360
Ala Leu Leu Gly Gly His Gln Gly Gly Ile Thr His Leu Cys Phe
               365                 370                 375
His Pro Asp Gly Asn Arg Phe Phe Ser Gly Ala Arg Lys Asp Ala
               380                 385                 390
Glu Leu Leu Cys Trp Asp Leu Arg Gln Ser Gly Tyr Pro Leu Trp
               395                 400                 405
```

Ser Leu Gly Arg Glu Val Thr Thr Asn Gln Arg Ile Tyr Phe Asp
410 415 420
Leu Asp Pro Thr Gly Gln Phe Leu Val Ser Gly Ser Thr Ser Gly
425 430 435
Ala Val Ser Val Trp Asp Thr Asp Gly Pro Gly Asn Asp Gly Lys
440 445 450
Pro Glu Pro Val Leu Ser Phe Leu Pro Gln Lys Asp Cys Thr Asn
455 460 465
Gly Val Ser Leu His Pro Ser Leu Pro Leu Leu Ala Thr Ala Ser
470 475 480
Gly Gln Arg Val Phe Pro Glu Pro Thr Glu Ser Gly Asp Glu Gly
485 490 495
Glu Glu Leu Gly Leu Pro Leu Leu Ser Thr Arg His Val His Leu
500 505 510
Glu Cys Arg Leu Gln Leu Trp Trp Cys Gly Gly Gly Pro Asp Ser
515 520 525
Ser Ile Pro Asp Asp His Gln Gly Glu Lys Gly Gln Gly Gly Thr
530 535 540
Gly Gly Arg Ser Trp Gly Ala
545

<210> 57
<211> 686
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3002667CD1

<400> 57
Met Gly Glu Phe Lys Val His Arg Val Arg Phe Phe Asn Tyr Val
1 5 10 15
Pro Ser Gly Ile Arg Cys Val Ala Tyr Asn Asn Gln Ser Asn Arg
20 25 30
Leu Ala Val Ser Arg Thr Asp Gly Thr Val Glu Ile Tyr Asn Leu
35 40 45
Ser Ala Asn Tyr Phe Gln Glu Lys Phe Phe Pro Gly His Glu Ser
50 55 60
Arg Ala Thr Glu Ala Leu Cys Trp Ala Glu Gly Gln Arg Leu Phe
65 70 75
Ser Ala Gly Leu Asn Gly Glu Ile Met Glu Tyr Asp Leu Gln Ala
80 85 90
Leu Asn Ile Lys Tyr Ala Met Asp Ala Phe Gly Gly Pro Ile Trp
95 100 105
Ser Met Ala Ala Ser Pro Ser Gly Ser Gln Leu Leu Val Gly Cys
110 115 120
Glu Asp Gly Ser Val Lys Leu Phe Gln Ile Thr Pro Asp Lys Ile
125 130 135
Gln Phe Glu Arg Asn Phe Asp Arg Gln Lys Ser Arg Ile Leu Ser
140 145 150
Leu Ser Trp His Pro Ser Gly Thr His Ile Ala Ala Gly Ser Ile
155 160 165
Asp Tyr Ile Ser Val Phe Asp Val Lys Ser Gly Ser Ala Val His
170 175 180
Lys Met Ile Val Asp Arg Gln Tyr Met Gly Val Ser Lys Arg Lys
185 190 195
Cys Ile Val Trp Gly Val Ala Phe Leu Ser Asp Gly Thr Ile Ile
200 205 210
Ser Val Asp Ser Ala Gly Lys Val Gln Phe Trp Asp Ser Ala Thr
215 220 225
Gly Thr Leu Val Lys Ser His Leu Ile Ala Asn Ala Asp Val Gln
230 235 240
Ser Ile Ala Val Ala Asp Gln Glu Asp Ser Phe Val Val Gly Thr

```
                    245                     250                     255
Ala Glu Gly Thr Val Phe His Phe Gln Leu Val Pro Val Thr Ser
                260                     265                     270
Asn Ser Ser Glu Lys Gln Trp Val Arg Thr Lys Pro Phe Gln His
            275                     280                     285
His Thr His Asp Val Arg Thr Val Ala His Ser Pro Thr Ala Leu
        290                     295                     300
Ile Ser Gly Gly Thr Asp Thr His Leu Val Phe Arg Pro Leu Met
305                     310                     315
Glu Lys Val Glu Val Lys Asn Tyr Asp Ala Ala Leu Arg Lys Ile
                320                     325                     330
Thr Phe Pro His Arg Cys Leu Ile Ser Cys Ser Lys Lys Arg Gln
            335                     340                     345
Leu Leu Leu Phe Gln Phe Ala His His Leu Glu Leu Trp Arg Leu
        350                     355                     360
Gly Ser Thr Val Ala Thr Gly Lys Asn Gly Asp Thr Leu Pro Leu
365                     370                     375
Ser Lys Asn Ala Asp His Leu Leu His Leu Lys Thr Lys Gly Pro
                380                     385                     390
Glu Asn Ile Ile Cys Ser Cys Ile Ser Pro Cys Gly Ser Trp Ile
            395                     400                     405
Ala Tyr Ser Thr Val Ser Arg Phe Phe Leu Tyr Arg Leu Asn Tyr
        410                     415                     420
Glu His Asp Asn Ile Ser Leu Lys Arg Val Ser Lys Met Pro Ala
425                     430                     435
Phe Leu Arg Ser Ala Leu Gln Ile Leu Phe Ser Glu Asp Ser Thr
                440                     445                     450
Lys Leu Phe Val Ala Ser Asn Gln Gly Ala Leu His Ile Val Gln
            455                     460                     465
Leu Ser Gly Gly Ser Phe Lys His Leu His Ala Phe Gln Pro Gln
        470                     475                     480
Ser Gly Thr Val Glu Ala Met Cys Leu Leu Ala Val Ser Pro Asp
485                     490                     495
Gly Asn Trp Leu Ala Ala Ser Gly Thr Ser Ala Gly Val His Val
                500                     505                     510
Tyr Asn Val Lys Gln Leu Lys Leu His Cys Thr Val Pro Ala Tyr
            515                     520                     525
Asn Phe Pro Val Thr Ala Met Ala Ile Ala Pro Asn Thr Asn Asn
        530                     535                     540
Leu Val Ile Ala His Ser Asp Gln Gln Val Phe Glu Tyr Ser Ile
545                     550                     555
Pro Asp Lys Gln Tyr Thr Asp Trp Ser Arg Thr Val Gln Lys Gln
                560                     565                     570
Gly Phe His His Leu Trp Leu Gln Arg Asp Thr Pro Ile Thr His
            575                     580                     585
Ile Ser Phe His Pro Lys Arg Pro Met His Ile Leu Leu His Asp
        590                     595                     600
Ala Tyr Met Phe Cys Ile Ile Asp Lys Ser Leu Pro Leu Pro Asn
605                     610                     615
Asp Lys Thr Leu Leu Tyr Asn Pro Phe Pro Pro Thr Asn Glu Ser
                620                     625                     630
Asp Val Ile Arg Arg Arg Thr Ala His Ala Phe Lys Ile Ser Lys
            635                     640                     645
Ile Tyr Lys Pro Leu Leu Phe Met Asp Leu Leu Asp Glu Arg Thr
        650                     655                     660
Leu Val Ala Val Glu Arg Pro Leu Asp Asp Ile Ile Ala Gln Leu
665                     670                     675
Pro Pro Pro Ile Lys Lys Lys Lys Phe Gly Thr
                680                     685
```

<210> 58
<211> 93
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3043734CD1

<400> 58
Met Thr Ser Lys Arg Lys Pro Cys Gln Thr Gln Leu Arg Arg Ser
1 5 10 15
Ile Ser Glu Gln Leu Arg Asp Ser Thr Ala Arg Ala Trp Asp Leu
20 25 30
Leu Trp Lys Asn Val Arg Glu Arg Arg Leu Ala Glu Ile Glu Ala
35 40 45
Lys Glu Ala Cys Asp Trp Leu Arg Ala Ala Gly Phe Pro Gln Tyr
50 55 60
Ala Gln Leu Tyr Glu Asp Ser Gln Phe Pro Ile Asn Ile Val Ala
65 70 75
Val Lys Asn Asp His Asp Phe Leu Glu Lys Asp Leu Val Glu Pro
80 85 90
Leu Cys Arg

<210> 59
<211> 521
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3294893CD1

<400> 59
Met Arg Arg Gly His Gly Gln Arg Arg Gly Gln Glu Ala Ile Leu
1 5 10 15
Glu Ala His Asn Ser Lys Leu Pro Gly Ser Ile Gln His Val Tyr
20 25 30
Gly Ala Gln His Pro Pro Phe Asp Pro Leu Leu His Gly Thr Leu
35 40 45
Leu Arg Ser Thr Ala Lys Met Pro Thr Thr Pro Val Lys Ala Lys
50 55 60
Arg Val Ser Thr Phe Gln Glu Phe Glu Ser Asn Thr Ser Asp Ala
65 70 75
Trp Asp Ala Gly Glu Asp Asp Asp Glu Leu Leu Ala Met Ala Ala
80 85 90
Glu Ser Leu Asn Ser Glu Val Val Met Glu Thr Ala Asn Arg Val
95 100 105
Leu Arg Asn His Ser Gln Arg Gln Gly Arg Pro Thr Leu Gln Glu
110 115 120
Gly Pro Gly Leu Gln Gln Lys Pro Arg Pro Glu Ala Glu Pro Pro
125 130 135
Ser Pro Pro Ser Gly Asp Leu Arg Leu Val Lys Ser Val Ser Glu
140 145 150
Ser His Thr Ser Cys Pro Ala Glu Ser Ala Ser Asp Ala Ala Pro
155 160 165
Leu Gln Arg Ser Gln Ser Leu Pro His Ser Ala Thr Val Thr Leu
170 175 180
Gly Gly Thr Ser Asp Pro Ser Thr Leu Ser Ser Ser Ala Leu Ser
185 190 195
Glu Arg Glu Ala Ser Arg Leu Asp Lys Phe Lys Gln Leu Leu Ala
200 205 210
Gly Pro Asn Thr Asp Leu Glu Glu Leu Arg Arg Leu Ser Trp Ser
215 220 225
Gly Ile Pro Lys Pro Val Arg Pro Met Thr Trp Lys Leu Leu Ser
230 235 240
Gly Tyr Leu Pro Ala Asn Val Asp Arg Arg Pro Ala Thr Leu Gln
245 250 255

Arg Lys Gln Lys Glu Tyr Phe Ala Phe Ile Glu His Tyr Tyr Asp
260 265 270
Ser Arg Asn Asp Glu Val His Gln Asp Thr Tyr Arg Gln Ile His
275 280 285
Ile Asp Ile Pro Arg Met Ser Pro Glu Ala Leu Ile Leu Gln Pro
290 295 300
Lys Val Thr Glu Ile Phe Glu Arg Ile Leu Phe Ile Trp Ala Ile
305 310 315
Arg His Pro Ala Ser Gly Tyr Val Gln Gly Ile Asn Asp Leu Val
320 325 330
Thr Pro Phe Phe Val Val Phe Ile Cys Glu Tyr Ile Glu Ala Glu
335 340 345
Glu Val Asp Thr Val Asp Val Ser Gly Val Pro Ala Glu Val Leu
350 355 360
Cys Asn Ile Glu Ala Asp Thr Tyr Trp Cys Met Ser Lys Leu Leu
365 370 375
Asp Gly Ile Gln Asp Asn Tyr Thr Phe Ala Gln Pro Gly Ile Gln
380 385 390
Met Lys Val Lys Met Leu Glu Glu Leu Val Ser Arg Ile Asp Glu
395 400 405
Gln Val His Arg His Leu Asp Gln His Glu Val Arg Tyr Leu Gln
410 415 420
Phe Ala Phe Arg Trp Met Asn Asn Leu Leu Met Arg Glu Val Pro
425 430 435
Leu Arg Cys Thr Ile Arg Leu Trp Asp Thr Tyr Gln Ser Glu Pro
440 445 450
Asp Gly Phe Ser His Phe His Leu Tyr Val Cys Ala Ala Phe Leu
455 460 465
Val Arg Trp Arg Lys Glu Ile Leu Glu Glu Lys Asp Phe Gln Glu
470 475 480
Leu Leu Leu Phe Leu Gln Asn Leu Pro Thr Ala His Trp Asp Asp
485 490 495
Glu Asp Ile Ser Leu Leu Leu Ala Glu Ala Tyr Arg Leu Lys Phe
500 505 510
Ala Phe Ala Asp Ala Pro Asn His Tyr Lys Lys
515 520

<210> 60
<211> 751
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3349052CD1

<400> 60
Met Arg Leu Leu Gly Ala Ala Ala Val Ala Ala Leu Gly Arg Gly
1 5 10 15
Arg Ala Pro Ala Ser Leu Gly Trp Gln Arg Lys Gln Val Asn Trp
20 25 30
Lys Ala Cys Arg Trp Ser Ser Ser Gly Val Ile Pro Asn Glu Lys
35 40 45
Ile Arg Asn Ile Gly Ile Ser Ala His Ile Asp Ser Gly Lys Thr
50 55 60
Thr Leu Thr Glu Arg Val Leu Tyr Tyr Thr Gly Arg Ile Ala Lys
65 70 75
Met His Glu Val Lys Gly Lys Asp Gly Val Gly Ala Val Met Asp
80 85 90
Ser Met Glu Leu Glu Arg Gln Arg Gly Ile Thr Ile Gln Ser Ala
95 100 105
Ala Thr Tyr Thr Met Trp Lys Asp Val Asn Ile Asn Ile Ile Asp
110 115 120
Thr Pro Gly His Val Asp Phe Thr Ile Glu Val Glu Arg Ala Leu

```
                125                 130                 135
Arg Val Leu Asp Gly Ala Val Leu Val Leu Cys Ala Val Gly Gly
                140                 145                 150
Val Gln Cys Gln Thr Met Thr Val Asn Arg Gln Met Lys Arg Tyr
                155                 160                 165
Asn Val Pro Phe Leu Thr Phe Ile Asn Lys Leu Asp Arg Met Gly
                170                 175                 180
Ser Asn Pro Ala Arg Ala Leu Gln Gln Met Arg Ser Lys Leu Asn
                185                 190                 195
His Asn Ala Ala Phe Met Gln Ile Pro Met Gly Leu Glu Gly Asn
                200                 205                 210
Phe Lys Gly Ile Ile Asp Leu Ile Glu Glu Arg Ala Ile Tyr Phe
                215                 220                 225
Asp Gly Asp Phe Gly Gln Ile Val Arg Tyr Gly Glu Ile Pro Ala
                230                 235                 240
Glu Leu Arg Ala Ala Ala Thr Asp His Arg Gln Glu Leu Ile Glu
                245                 250                 255
Cys Val Ala Asn Ser Asp Glu Gln Leu Gly Glu Met Phe Leu Glu
                260                 265                 270
Glu Lys Ile Pro Ser Ile Ser Asp Leu Lys Leu Ala Ile Arg Arg
                275                 280                 285
Ala Thr Leu Lys Arg Ser Phe Thr Pro Val Phe Leu Gly Ser Ala
                290                 295                 300
Leu Lys Asn Lys Gly Val Gln Pro Leu Leu Asp Ala Val Leu Glu
                305                 310                 315
Tyr Leu Pro Asn Pro Ser Glu Val Gln Asn Tyr Ala Ile Leu Asn
                320                 325                 330
Lys Glu Asp Asp Ser Lys Glu Lys Thr Lys Ile Leu Met Asn Ser
                335                 340                 345
Ser Arg Asp Asn Ser His Pro Phe Val Gly Leu Ala Phe Lys Leu
                350                 355                 360
Glu Val Gly Arg Phe Gly Gln Leu Thr Tyr Val Arg Ser Tyr Gln
                365                 370                 375
Gly Glu Leu Lys Lys Gly Asp Thr Ile Tyr Asn Thr Arg Thr Arg
                380                 385                 390
Lys Lys Val Arg Leu Gln Arg Leu Ala Arg Met His Ala Asp Met
                395                 400                 405
Met Glu Asp Val Glu Glu Val Tyr Ala Gly Asp Ile Cys Ala Leu
                410                 415                 420
Phe Gly Ile Asp Cys Ala Ser Gly Asp Thr Phe Thr Asp Lys Ala
                425                 430                 435
Asn Ser Gly Leu Ser Met Glu Ser Ile His Val Pro Asp Pro Val
                440                 445                 450
Ile Ser Ile Ala Met Lys Pro Ser Asn Lys Asn Asp Leu Glu Lys
                455                 460                 465
Phe Ser Lys Gly Ile Gly Arg Phe Thr Arg Glu Asp Pro Thr Phe
                470                 475                 480
Lys Val Tyr Phe Asp Thr Glu Asn Lys Glu Thr Val Ile Ser Gly
                485                 490                 495
Met Gly Glu Leu His Leu Glu Ile Tyr Ala Gln Arg Leu Glu Arg
                500                 505                 510
Glu Tyr Gly Cys Pro Cys Ile Thr Gly Lys Pro Lys Val Ala Phe
                515                 520                 525
Arg Glu Thr Ile Thr Ala Pro Val Pro Phe Asp Phe Thr His Lys
                530                 535                 540
Lys Gln Ser Gly Gly Ala Gly Gln Tyr Gly Lys Val Ile Gly Val
                545                 550                 555
Leu Glu Pro Leu Asp Pro Glu Asp Tyr Thr Lys Leu Glu Phe Ser
                560                 565                 570
Asp Glu Thr Phe Gly Ser Asn Ile Pro Lys Gln Phe Val Pro Ala
                575                 580                 585
Val Glu Lys Gly Phe Leu Asp Ala Cys Glu Lys Gly Pro Leu Ser
                590                 595                 600
```

```
Gly His Lys Leu Ser Gly Leu Arg Phe Val Leu Gln Asp Gly Ala
                605                 610                 615
His His Met Val Asp Ser Asn Glu Ile Ser Phe Ile Arg Ala Gly
                620                 625                 630
Glu Gly Ala Leu Lys Gln Ala Leu Ala Asn Ala Thr Leu Cys Ile
                635                 640                 645
Leu Glu Pro Ile Met Ala Val Glu Val Val Ala Pro Asn Glu Phe
                650                 655                 660
Gln Gly Gln Val Ile Ala Gly Ile Asn Arg Arg His Gly Val Ile
                665                 670                 675
Thr Gly Gln Asp Gly Val Glu Asp Tyr Phe Thr Leu Tyr Ala Asp
                680                 685                 690
Val Pro Leu Asn Asp Met Phe Gly Tyr Ser Thr Glu Leu Arg Ser
                695                 700                 705
Cys Thr Glu Gly Lys Gly Glu Tyr Thr Met Glu Tyr Ser Arg Tyr
                710                 715                 720
Gln Pro Cys Leu Pro Ser Thr Gln Glu Asp Val Ile Asn Lys Tyr
                725                 730                 735
Leu Glu Ala Thr Gly Gln Leu Pro Val Lys Lys Gly Lys Ala Lys
                740                 745                 750
Asn
```

<210> 61
<211> 666
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3357264CD1

<220>
<221> unsure
<222> 281
<223> unknown or other

<400> 61

```
Met Cys Gly Ala Val Ile Pro Leu His Lys Pro Ala Gly Arg Lys
 1               5                  10                  15
Leu Gln Asn Gln Arg Ala Ala Leu Asn Gln Gln Ile Leu Lys Ala
                20                  25                  30
Val Arg Met Arg Thr Gly Ala Glu Asn Leu Leu Lys Val Ala Thr
                35                  40                  45
Asn Ser Lys Val Arg Glu Gln Val Arg Leu Glu Leu Ser Phe Val
                50                  55                  60
Asn Ser Asp Leu Gln Met Leu Lys Glu Glu Leu Glu Gly Leu Asn
                65                  70                  75
Ile Ser Val Gly Val Tyr Gln Asn Thr Glu Glu Ala Phe Thr Ile
                80                  85                  90
Pro Leu Ile Pro Leu Gly Leu Lys Glu Thr Lys Asp Val Asp Phe
                95                 100                 105
Ala Val Val Leu Lys Asp Phe Ile Leu Glu His Tyr Ser Glu Asp
               110                 115                 120
Gly Tyr Leu Tyr Glu Asp Glu Ile Ala Asp Leu Met Asp Leu Arg
               125                 130                 135
Gln Ala Cys Arg Thr Pro Ser Arg Asp Glu Ala Gly Val Glu Leu
               140                 145                 150
Leu Met Thr Tyr Phe Ile Gln Leu Gly Phe Val Glu Ser Arg Phe
               155                 160                 165
Phe Pro Pro Thr Arg Gln Met Gly Leu Leu Phe Thr Trp Tyr Asp
               170                 175                 180
Ser Leu Thr Gly Val Pro Val Ser Gln Gln Asn Leu Leu Leu Glu
               185                 190                 195
```

Lys Ala Ser Val Leu Phe Asn Thr Gly Ala Leu Tyr Thr Gln Ile
200 205 210
Gly Thr Arg Cys Asp Arg Gln Thr Gln Ala Gly Leu Glu Ser Ala
215 220 225
Ile Asp Ala Phe Gln Arg Ala Ala Gly Val Leu Asn Tyr Leu Lys
230 235 240
Asp Thr Phe Thr His Thr Pro Ser Tyr Asp Met Ser Pro Ala Met
245 250 255
Leu Ser Val Leu Val Lys Met Met Leu Ala Gln Ala Gln Glu Ser
260 265 270
Val Phe Glu Lys Ile Ser Leu Pro Gly Ile Xaa Asn Glu Phe Phe
275 280 285
Met Leu Val Lys Val Ala Gln Glu Ala Ala Lys Val Gly Glu Val
290 295 300
Tyr Gln Gln Leu His Ala Ala Met Ser Gln Ala Pro Val Lys Glu
305 310 315
Asn Ile Pro Tyr Ser Trp Ala Ser Leu Ala Cys Val Lys Ala His
320 325 330
His Tyr Ala Ala Leu Ala His Tyr Phe Thr Ala Ile Leu Leu Ile
335 340 345
Asp His Gln Val Lys Pro Gly Thr Asp Leu Asp His Gln Glu Lys
350 355 360
Cys Leu Ser Gln Leu Tyr Asp His Met Pro Glu Gly Leu Thr Pro
365 370 375
Leu Ala Thr Leu Lys Asn Asp Gln Gln Arg Arg Gln Leu Gly Lys
380 385 390
Ser His Leu Arg Arg Ala Met Ala His His Glu Glu Ser Val Arg
395 400 405
Glu Ala Ser Leu Cys Lys Lys Leu Arg Thr Ile Glu Val Leu Gln
410 415 420
Lys Val Leu Cys Ala Ala Gln Glu Arg Ser Arg Leu Thr Tyr Ala
425 430 435
Gln His Gln Glu Glu Asp Asp Leu Leu Asn Leu Ile Asp Ala Pro
440 445 450
Ser Val Val Ala Lys Thr Glu Gln Glu Val Asp Ile Ile Leu Pro
455 460 465
Gln Phe Ser Lys Leu Thr Val Thr Asp Phe Phe Gln Lys Leu Gly
470 475 480
Pro Leu Ser Val Phe Ser Ala Asn Lys Arg Trp Thr Pro Pro Arg
485 490 495
Ser Ile Arg Phe Thr Ala Glu Glu Gly Asp Leu Gly Phe Thr Leu
500 505 510
Arg Gly Asn Ala Pro Val Gln Val His Phe Leu Asp Pro Tyr Cys
515 520 525
Ser Ala Ser Val Ala Gly Ala Arg Glu Gly Asp Tyr Ile Val Ser
530 535 540
Ile Gln Leu Val Asp Cys Lys Trp Leu Thr Leu Ser Glu Val Met
545 550 555
Lys Leu Leu Lys Ser Phe Gly Glu Asp Glu Ile Glu Met Lys Val
560 565 570
Val Ser Leu Leu Asp Ser Thr Ser Ser Met His Asn Lys Ser Ala
575 580 585
Thr Tyr Ser Val Gly Met Gln Lys Thr Tyr Ser Met Ile Cys Leu
590 595 600
Ala Ile Asp Asp Asp Asp Lys Thr Asp Lys Thr Lys Lys Ile Ser
605 610 615
Lys Lys Leu Ser Phe Leu Ser Trp Gly Thr Asn Lys Asn Arg Gln
620 625 630
Lys Ser Ala Ser Thr Leu Cys Leu Pro Ser Val Gly Ala Ala Arg
635 640 645
Pro Gln Val Lys Lys Lys Leu Pro Ser Pro Phe Ser Leu Leu Asn
650 655 660
Ser Asp Ser Ser Trp Tyr

665

<210> 62
<211> 746
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3576329CD1

<400> 62
Met Ala Gly Ser Arg Gly Ala Gly Arg Thr Ala Ala Pro Ser Val
1 5 10 15
Arg Pro Glu Lys Arg Arg Ser Glu Pro Glu Leu Glu Pro Glu Pro
20 25 30
Glu Pro Glu Pro Pro Leu Leu Cys Thr Ser Pro Leu Ser His Ser
35 40 45
Thr Gly Ser Asp Ser Gly Val Ser Asp Ser Glu Glu Ser Val Phe
50 55 60
Ser Gly Leu Glu Asp Ser Gly Ser Asp Ser Ser Glu Asp Asp Asp
65 70 75
Glu Gly Asp Glu Glu Gly Glu Asp Gly Ala Leu Asp Asp Glu Gly
80 85 90
His Ser Gly Ile Lys Lys Thr Thr Glu Glu Gln Val Gln Ala Ser
95 100 105
Thr Pro Cys Pro Arg Thr Glu Met Ala Ser Ala Arg Ile Gly Asp
110 115 120
Glu Tyr Ala Glu Asp Ser Ser Asp Glu Glu Asp Ile Arg Asn Thr
125 130 135
Val Gly Asn Val Pro Leu Glu Trp Tyr Asp Asp Phe Pro His Val
140 145 150
Gly Tyr Asp Leu Asp Gly Arg Arg Ile Tyr Lys Pro Leu Arg Thr
155 160 165
Arg Asp Glu Leu Asp Gln Phe Leu Asp Lys Met Asp Asp Pro Asp
170 175 180
Tyr Trp Arg Thr Val Gln Asp Pro Met Thr Gly Arg Asp Leu Arg
185 190 195
Leu Thr Asp Glu Gln Val Ala Leu Val Arg Arg Leu Gln Ser Gly
200 205 210
Gln Phe Gly Asp Val Gly Phe Asn Pro Tyr Glu Pro Ala Val Asp
215 220 225
Phe Phe Ser Gly Asp Val Met Ile His Pro Val Thr Asn Arg Pro
230 235 240
Ala Asp Lys Arg Ser Phe Ile Pro Ser Leu Val Glu Lys Glu Lys
245 250 255
Val Ser Arg Met Val His Ala Ile Lys Met Gly Trp Ile Gln Pro
260 265 270
Arg Arg Pro Arg Asp Pro Thr Pro Ser Phe Tyr Asp Leu Trp Ala
275 280 285
Gln Glu Asp Pro Asn Ala Val Leu Gly Arg His Lys Met His Val
290 295 300
Pro Ala Pro Lys Leu Ala Leu Pro Gly His Ala Glu Ser Tyr Asn
305 310 315
Pro Pro Pro Glu Tyr Leu Leu Ser Glu Glu Glu Arg Leu Ala Trp
320 325 330
Glu Gln Gln Glu Pro Gly Glu Arg Lys Leu Gly Phe Leu Pro Arg
335 340 345
Lys Phe Pro Ser Leu Arg Ala Val Pro Ala Tyr Gly Arg Phe Ile
350 355 360
Gln Glu Arg Phe Glu Arg Cys Leu Asp Leu Tyr Leu Cys Pro Arg
365 370 375
Gln Arg Lys Met Arg Val Asn Val Asp Pro Glu Asp Leu Ile Pro
380 385 390

Lys Leu Pro Arg Pro Arg Asp Leu Gln Pro Phe Pro Thr Cys Gln
395 400 405
Ala Leu Val Tyr Arg Gly His Ser Asp Leu Val Arg Cys Leu Ser
410 415 420
Val Ser Pro Gly Gly Gln Trp Leu Val Ser Gly Ser Asp Asp Gly
425 430 435
Ser Leu Arg Leu Trp Glu Val Ala Thr Ala Arg Cys Val Arg Thr
440 445 450
Val Pro Val Gly Gly Val Val Lys Ser Val Ala Trp Asn Pro Ser
455 460 465
Pro Ala Val Cys Leu Val Ala Ala Ala Val Glu Asp Ser Val Leu
470 475 480
Leu Leu Asn Pro Ala Leu Gly Asp Arg Leu Val Ala Gly Ser Thr
485 490 495
Asp Gln Leu Leu Ser Ala Phe Val Pro Pro Glu Glu Pro Pro Leu
500 505 510
Gln Pro Ala Arg Trp Leu Glu Ala Ser Glu Glu Glu Arg Gln Val
515 520 525
Gly Leu Arg Leu Arg Ile Cys His Gly Lys Pro Val Thr Gln Val
530 535 540
Thr Trp His Gly Arg Gly Asp Tyr Leu Ala Val Val Leu Ala Thr
545 550 555
Gln Gly His Thr Gln Val Leu Ile His Gln Leu Ser Arg Arg Arg
560 565 570
Ser Gln Ser Pro Phe Arg Arg Ser His Gly Gln Val Gln Arg Val
575 580 585
Ala Phe His Pro Ala Arg Pro Phe Leu Leu Val Ala Ser Gln Arg
590 595 600
Ser Val Arg Leu Tyr His Leu Leu Arg Gln Glu Leu Thr Lys Lys
605 610 615
Leu Met Pro Asn Cys Lys Trp Val Ser Ser Leu Ala Val His Pro
620 625 630
Ala Gly Asp Asn Val Ile Cys Gly Ser Tyr Asp Ser Lys Leu Val
635 640 645
Trp Phe Asp Leu Asp Leu Ser Thr Lys Pro Tyr Arg Met Leu Arg
650 655 660
His His Lys Lys Ala Leu Arg Ala Val Ala Phe His Pro Arg Tyr
665 670 675
Pro Leu Phe Ala Ser Gly Ser Asp Asp Gly Ser Val Ile Val Cys
680 685 690
His Gly Met Val Tyr Asn Asp Leu Leu Gln Asn Pro Leu Leu Val
695 700 705
Pro Val Lys Val Leu Lys Gly His Val Leu Thr Arg Asp Leu Gly
710 715 720
Val Leu Asp Val Ile Phe His Pro Thr Gln Pro Trp Val Phe Ser
725 730 735
Ser Gly Ala Asp Gly Thr Val Arg Leu Phe Thr
740 745

<210> 63
<211> 212
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3805550CD1

<400> 63
Met Ala Gly Pro Gly Pro Gly Pro Gly Asp Pro Asp Glu Gln Tyr
1 5 10 15
Asp Phe Leu Phe Lys Leu Val Leu Val Gly Asp Ala Ser Val Gly
20 25 30
Lys Thr Cys Val Val Gln Arg Phe Lys Thr Gly Ala Phe Ser Glu

```
                35                      40                      45
Arg Gln Gly Ser Thr Ile Gly Val Asp Phe Thr Met Lys Thr Leu
                50                      55                      60
Glu Ile Gln Gly Lys Arg Val Lys Leu Gln Ile Trp Asp Thr Ala
                65                      70                      75
Gly Gln Glu Arg Phe Arg Thr Ile Thr Gln Ser Tyr Tyr Arg Ser
                80                      85                      90
Ala Asn Gly Ala Ile Leu Ala Tyr Asp Ile Thr Lys Arg Ser Ser
                95                      100                     105
Phe Leu Ser Val Pro His Trp Ile Glu Asp Val Arg Lys Tyr Ala
                110                     115                     120
Gly Ser Asn Ile Val Gln Leu Leu Ile Gly Asn Lys Ser Asp Leu
                125                     130                     135
Ser Glu Leu Arg Glu Val Ser Leu Ala Glu Ala Gln Ser Leu Ala
                140                     145                     150
Glu His Tyr Asp Ile Leu Cys Ala Ile Glu Thr Ser Ala Lys Asp
                155                     160                     165
Ser Ser Asn Val Glu Glu Ala Phe Leu Arg Val Ala Thr Glu Leu
                170                     175                     180
Ile Met Arg His Gly Gly Pro Leu Phe Ser Glu Lys Ser Pro Asp
                185                     190                     195
His Ile Gln Leu Asn Ser Lys Asp Ile Gly Glu Gly Trp Gly Cys
                200                     205                     210
Gly Cys

<210> 64
<211> 307
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4546403CD1

<400> 64
Met Arg Cys Leu His Ser Glu Lys Ala His Asp Leu Gly Ile Thr
  1               5                       10                      15
Cys Cys Asp Phe Ser Ser Gln Pro Val Ser Asp Gly Glu Gln Gly
                20                      25                      30
Leu Gln Phe Phe Arg Leu Ala Ser Cys Gly Gln Asp Cys Gln Val
                35                      40                      45
Lys Ile Trp Ile Val Ser Phe Thr His Ile Leu Gly Phe Glu Leu
                50                      55                      60
Lys Tyr Lys Ser Thr Leu Ser Gly His Cys Ala Pro Val Leu Ala
                65                      70                      75
Cys Ala Phe Ser His Asp Gly Gln Met Leu Val Ser Gly Ser Val
                80                      85                      90
Asp Lys Ser Val Ile Val Tyr Asp Thr Asn Thr Glu Asn Ile Leu
                95                      100                     105
His Thr Leu Thr Gln His Thr Arg Tyr Val Thr Thr Cys Ala Phe
                110                     115                     120
Ala Pro Asn Thr Leu Leu Leu Ala Thr Gly Ser Met Asp Lys Thr
                125                     130                     135
Val Asn Ile Trp Gln Phe Asp Leu Glu Thr Leu Cys Gln Ala Arg
                140                     145                     150
Ser Thr Glu His Gln Leu Lys Gln Phe Thr Glu Asp Trp Ser Glu
                155                     160                     165
Glu Asp Val Ser Thr Trp Leu Cys Ala Gln Asp Leu Lys Asp Leu
                170                     175                     180
Val Gly Ile Phe Lys Met Asn Asn Ile Asp Gly Lys Glu Leu Leu
                185                     190                     195
Asn Leu Thr Lys Glu Ser Leu Ala Asp Asp Leu Lys Ile Glu Ser
                200                     205                     210
```

```
Leu Gly Leu Arg Ser Lys Val Leu Arg Lys Ile Glu Glu Leu Arg
                215                 220                 225
Thr Lys Val Lys Ser Leu Ser Ser Gly Ile Pro Asp Glu Phe Ile
                230                 235                 240
Cys Pro Ile Thr Arg Glu Leu Met Lys Asp Pro Val Ile Ala Ser
                245                 250                 255
Asp Gly Tyr Ser Tyr Glu Lys Glu Ala Met Glu Asn Trp Ile Ser
                260                 265                 270
Lys Lys Lys Arg Thr Ser Pro Met Thr Asn Leu Val Leu Pro Ser
                275                 280                 285
Ala Val Leu Thr Pro Asn Arg Thr Leu Lys Met Ala Ile Asn Arg
                290                 295                 300
Trp Leu Glu Thr His Gln Lys
                305
```

<210> 65
<211> 378
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4767318CD1

<400> 65

```
Met Arg Ala Ala Ala Ala Pro Gly Leu Thr Ala Pro Trp Arg Leu
  1               5                  10                  15
Leu Gln Cys Cys Glu Leu Glu Ala Gly Glu Leu Gly Met Ala Val
                 20                  25                  30
Pro Ala Ala Ala Met Gly Pro Ser Ala Leu Gly Gln Ser Gly Pro
                 35                  40                  45
Gly Ser Met Ala Pro Trp Cys Ser Val Ser Ser Gly Pro Ser Arg
                 50                  55                  60
Tyr Val Leu Gly Met Gln Glu Leu Phe Arg Gly His Ser Lys Thr
                 65                  70                  75
Arg Glu Phe Leu Ala His Ser Ala Lys Val His Ser Val Ala Trp
                 80                  85                  90
Ser Cys Asp Gly Arg Arg Leu Ala Ser Gly Ser Phe Asp Lys Thr
                 95                 100                 105
Ala Ser Val Phe Leu Leu Glu Lys Asp Arg Leu Val Lys Glu Asn
                110                 115                 120
Asn Tyr Arg Gly His Gly Asp Ser Val Asp Gln Leu Cys Trp His
                125                 130                 135
Pro Ser Asn Pro Asp Leu Phe Val Thr Ala Ser Gly Asp Lys Thr
                140                 145                 150
Ile Arg Ile Trp Asp Val Arg Thr Thr Lys Cys Ile Ala Thr Val
                155                 160                 165
Asn Thr Lys Gly Glu Asn Ile Asn Ile Cys Trp Ser Pro Asp Gly
                170                 175                 180
Gln Thr Ile Ala Val Gly Asn Lys Asp Asp Val Val Thr Phe Ile
                185                 190                 195
Asp Ala Lys Thr His Arg Ser Lys Ala Glu Glu Gln Phe Lys Phe
                200                 205                 210
Glu Val Asn Glu Ile Ser Trp Asn Asn Asp Asn Asn Met Phe Phe
                215                 220                 225
Leu Thr Asn Gly Asn Gly Cys Ile Asn Ile Leu Ser Tyr Pro Glu
                230                 235                 240
Leu Lys Pro Val Gln Ser Ile Asn Ala His Pro Ser Asn Cys Ile
                245                 250                 255
Cys Ile Lys Phe Asp Pro Met Gly Lys Tyr Phe Ala Thr Gly Ser
                260                 265                 270
Ala Asp Ala Leu Val Ser Leu Trp Asp Val Asp Glu Leu Val Cys
                275                 280                 285
Val Arg Cys Phe Ser Arg Leu Asp Trp Pro Val Arg Thr Leu Ser
```

```
                290                     295                     300
Phe Ser His Asp Gly Lys Met Leu Ala Ser Ala Ser Glu Asp His
                305                     310                     315
Phe Ile Asp Ile Ala Glu Val Glu Thr Gly Asp Lys Leu Trp Glu
                320                     325                     330
Val Gln Cys Glu Ser Pro Thr Phe Thr Val Ala Trp His Pro Lys
                335                     340                     345
Arg Pro Leu Leu Ala Phe Ala Cys Asp Asp Lys Asp Gly Lys Tyr
                350                     355                     360
Asp Ser Ser Arg Glu Ala Gly Thr Val Lys Leu Phe Gly Leu Pro
                365                     370                     375
Asn Asp Ser

<210> 66
<211> 466
<212> PRT
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4834527CD1

<400> 66
Met Pro Gln Thr Leu Ser Ala Ser Asp Met Val Thr Pro Gly Ser
  1               5                  10                  15
Leu Ser Pro Pro Pro Thr Glu Pro Thr Asp Gly Glu Gln Ala Gly
                 20                  25                  30
Gln Pro Leu Leu Asp Gly Ala Pro Ser Ser Ala Ser Leu Glu Thr
                 35                  40                  45
Leu Ile Gln His Leu Val Pro Thr Ala Asp Tyr Tyr Pro Glu Lys
                 50                  55                  60
Ala Tyr Ile Phe Thr Phe Leu Leu Ser Ser Arg Leu Phe Ile Glu
                 65                  70                  75
Pro Arg Glu Leu Leu Ala Arg Val Cys His Leu Cys Ile Glu Gln
                 80                  85                  90
Gln Gln Leu Asp Lys Pro Val Leu Asp Lys Ala Arg Val Arg Lys
                 95                 100                 105
Phe Gly Pro Lys Leu Leu Gln Leu Leu Ala Glu Trp Thr Glu Thr
                110                 115                 120
Phe Pro Arg Asp Phe Gln Glu Glu Ser Thr Ile Gly His Leu Lys
                125                 130                 135
Asp Val Val Gly Arg Ile Ala Pro Cys Asp Glu Ala Tyr Arg Lys
                140                 145                 150
Arg Met His Gln Leu Leu Gln Ala Leu His Gln Lys Leu Ala Ala
                155                 160                 165
Leu Arg Gln Gly Pro Glu Gly Leu Val Gly Ala Asp Lys Pro Ile
                170                 175                 180
Ser Tyr Arg Thr Lys Pro Pro Ala Ser Ile His Arg Glu Leu Leu
                185                 190                 195
Gly Val Cys Ser Asp Pro Tyr Thr Leu Ala Gln Gln Leu Thr His
                200                 205                 210
Val Glu Leu Glu Arg Leu Arg His Ile Gly Pro Glu Glu Phe Val
                215                 220                 225
Gln Ala Phe Val Asn Lys Asp Pro Leu Ala Ser Thr Lys Pro Cys
                230                 235                 240
Phe Ser Asp Lys Thr Ser Asn Leu Glu Ala Tyr Val Lys Trp Phe
                245                 250                 255
Asn Arg Leu Cys Tyr Leu Val Ala Thr Glu Ile Cys Met Pro Ala
                260                 265                 270
Lys Lys Lys Gln Arg Ala Gln Val Ile Glu Phe Phe Ile Asp Val
                275                 280                 285
Ala Arg Glu Cys Phe Asn Ile Gly Asn Phe Asn Ser Leu Met Ala
                290                 295                 300
```

```
Ile Ile Ser Gly Met Asn Met Ser Pro Val Ser Arg Leu Lys Lys
                305                 310                 315
Thr Trp Ala Lys Val Arg Thr Ala Lys Phe Phe Ile Leu Glu His
                320                 325                 330
Gln Met Asp Pro Thr Gly Asn Phe Cys Asn Tyr Arg Thr Ala Leu
                335                 340                 345
Arg Gly Ala Ala His Arg Ser Leu Thr Ala His Ser Ser Arg Glu
                350                 355                 360
Lys Ile Val Ile Pro Phe Phe Ser Leu Leu Ile Lys Asp Ile Tyr
                365                 370                 375
Phe Leu Asn Glu Gly Cys Ala Asn Arg Leu Pro Asn Gly His Val
                380                 385                 390
Asn Phe Glu Lys Phe Leu Glu Leu Ala Lys Gln Val Gly Glu Phe
                395                 400                 405
Ile Thr Trp Lys Gln Val Glu Cys Pro Phe Glu Gln Asp Ala Ser
                410                 415                 420
Ile Thr His Tyr Leu Tyr Thr Ala Pro Ile Phe Ser Glu Asp Gly
                425                 430                 435
Leu Tyr Leu Ala Ser Tyr Glu Ser Glu Ser Pro Glu Asn Gln Thr
                440                 445                 450
Glu Lys Glu Arg Trp Lys Ala Leu Arg Ser Ser Ile Leu Gly Lys
                455                 460                 465
Thr
```

<210> 67
<211> 891
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1405545CB1

<400> 67

```
ggagaatggc ggcgccgggc tgcggctggg agcgggaaga ctctttgaaa tgcctgcggt 60
gctagagcga ctgagccgct ataatagcac gtcccaagct tttgctgagg tgctgcggct 120
gccgaagcag cagctgagga agctgctgta cccgctgcag gaagtagagc ggttcctcgc 180
cccctacggg aggcaagacc ttcacctgcg tatctttgac ccaagcccgg aggacatagc 240
cagggcggac aacatcttca cggccactga acggaaccgc atcgactacg tcagctccgc 300
cgtccgtatc gaccacgccc cggaccttcc gcggccagag gtgtgtttta taggcagaag 360
caatgttgga aaatcatctc taatcaaggc tttattttca ctggcccctg aggttgaagt 420
cagagtctcc aaaaaaccag gacacacaaa gaaaatgaat tttttcaaag ttggaaaaca 480
ttttacagtg gtggacatgc caggttatgg ctttagagca cctgaagatt ttgttgacat 540
ggtagagacc tatctaaaag aacgaaggaa cttgaagaga acatttttat tagtggatag 600
cgttgttgga attcaaaaaa cagacaatat tgccatagaa atgtgtgaag aatttgcatt 660
accttatgtg attgtattaa caaaaattga caaatcttcc aagggacatc ttttaaaaca 720
agtgcttcag atccagaaat ttgttaacat gaaaactcaa ggatgttttc ctcagttgtt 780
tcctgtaagt gctgtgacct ttctggaat ccacctgttg agatgcttta tagccagtgt 840
aacaggaagt cttgactaat ggttcccggt ttagctgaag attcaaaaaa a          891
```

<210> 68
<211> 1512
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1451265CB1

<400> 68

```
gcccatggag gtggccgtgt gtacggactc ggcggccccg atgtggagct gcatcgtgtg 60
ggaacttcac tcgggcgcca acctgctcac ctaccgcggc ggccaggcgg gaccccgcgg 120
cctggcgctg ctcaatggcg agtatctgct ggcggcgcag ctgggcaaga attacatcag 180
```

cgcctgggag ctccagcgga aggaccagct ccagcagaag atcatgtgcc ccgggcctgt 240
cacctgtctg actgcatcac ccaatggtct ctacgtcctg gcaggagttg cagaaagcat 300
ccacctgtgg gaggtctcca ccgggaacct tctggtcatc ctgagtcgac actaccagga 360
cgtctcctgc cttcagttca caggggacag cagccacttc atctcagggg gcaaggactg 420
cctggtgctg gtttggagcc tctgcagcgt gctgcaggcc gacccctcca ggattccggc 480
gcccaggcac gtctggtctc accacacgct ccccatcacg gacctgcact gcggctttgg 540
gggcccctg gcccgggtgg ccacctcctc actggaccag acggtgaagc tatgggaggt 600
ctcctcgggg gagctgctgc tctccgtcct ctttgacgtg tccatcatgg cagtgaccat 660
ggacctggct gagcaccata tgttctgcgg gggcagtgag ggctccatct tccaggtcga 720
cctcttcacc tggcccggac agagggagag gagcttccac ccagagcagg acgccgggaa 780
ggtcttcaaa gggcacagga accaggtgac ttgcctgtca gtgtccactg acggcagcgt 840
gctgctctca ggctcccacg acgagaccgt gcgcctctgg gacgtgcaga gcaagcagtg 900
catccggacg gtggccctca aaggcccagt caccaatgcc gccatcctgc tggcgcccgt 960
cagcatgctg agctcagact tcaggcccag cctgccgctg ccccacttca acaagcacct 1020
gctgggcgcc gagcacgggg acgagccgcg ccacgggggc ctcactctgc gcctgggcct 1080
ccaccagcag ggctcggagc ccagctacct ggaccgcacg gagcagctgc aggccgtcct 1140
gtgcagcacc atggagaaga gcgtgctcgg cggccaggac cagctgcgcg tccgtgtgac 1200
ggagctggag gacgaggtgc gcaacctgcg caagatcaat cgggacctgt tcgacttctc 1260
cacgcgcttc atcacgcggc cggccaagtg aggcccggag accccggccc gaggcgccca 1320
ggcctgagcc ccatgcctcc cagcaaccag ggcccgcggg tgtggccccc accagcccag 1380
gcctggactc tcctcagttc tgtgtcgtgt tcgggttttt cctctgtgac tgggccgtct 1440
tggtgtctcg tggcacgcgt cacagtggtg ctagtctgtt tttaacaaaa gaggatgaaa 1500
aaaaaaaaaa aa 1512

<210> 69
<211> 2536
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1556311CB1

<400> 69
caactcttgt tgaagctttt aggcgtcgca gactcttcat ttgtgagggc gacctctccc 60
gaggggctct tttcacacaa atatccccac ggcgtttctc ggaggcaccc ccgtcatacg 120
tcttgtctct cgcgacaatt ctctttgaag gcgaggcatt tcaccacaac tcttttcaac 180
caacctggcg acaacaccca gagcttacat tgaccccaat ttgaattttc atcggcccaa 240
ggctttcttt acactcaggg aactctcaca ctccttaggg ggaaaaaggc ttcgttaagg 300
gccttgcaag ggttaccggg ttccggaatt ttcccggggg cccctcggct ggccaggact 360
gaaacccaga cgagcatgcc agaaacagtc aaccataaca aacatgggaa cgtagctctc 420
cctggaacga aaccaactcc catccctcca ccccggctga agaagcaggc ttcttttctg 480
gaagcagagg gcggtgcaaa gaccttgagc ggcggccggc cgggcgcagg cccggagctg 540
gagctgggca cagctggcag cccaggtggg gccccgcctg aggccgcccc gggggattgc 600
acaagggccc cgccgcccag ctctgaatca cggcccccgt gccatggagg ccggcagcgg 660
ctgagcgaca tgagcatttc tacttcctcc tccgactcgc tggagttcga ccggagcatg 720
cctctgtttg gctacgaggc ggacaccaac agcagcctgg aggactacga gggggaaagt 780
gaccaagaga ccatggcgcc ccccatcaag tccaaaaaga aaaggagcag ctccttcgtg 840
ctgcccaagc tcgtcaagtc ccagctgcag aaggtgagcg gggtgttcag ctccttcatg 900
accccggaga agcggatggt ccgcaggatc gccgagcttt cccgggacaa atgcacctac 960
ttcgggtgct tagtgcagga ctacgtgagc ttcctgcagg agaacaagga gtgccacgtg 1020
tccagcaccg acatgctgca gaccatccgg cagttcatga cccaggtcaa gaactatttg 1080
tctcagagct cggagctgga cccccccatc gagtcgctga tccctgaaga ccaaatagat 1140
gtggtgctgg aaaaagccat gcacaagtgc atcttgaagc cctcaaggg gcacgtggag 1200
gccatgctga aggactttca catggccgat ggctcatgga agcaactcaa ggagaacctg 1260
cagcttgtgc ggcagaggaa tccgcaggag ctgggggtct tcgccccgac ccctgatttt 1320
gtggatgtgg agaaaatcaa agtcaagttc atgaccatgc agaagatgta ttcgccggaa 1380
aagaaggtca tgctgctgct gcgggtctgc aagctcattt acacggtcat ggagaacaac 1440
tcagggagga tgtatggcgc tgatgacttc ttgccagtcc tgacctatgt catagcccag 1500
tgtgacatgc ttgaattgga cactgaaatc gagtacatga tggagctcct agacccatcg 1560
ctgttacatg gagaaggagg ctattacttg acaagcgcat atggagcact ttctctgata 1620
aagaatttcc aagaagaaca agcagcgcga ctgctcagct cagaaaccag agacaccctg 1680
aggcagtggc acaaacggag aaccaccaac cggaccatcc cctctgtgga cgacttccag 1740

aattacctcc gagttgcatt tcaggaggtc aacagtggtt gcacaggaaa gaccctcctt 1800
gtgagacctt acatcaccac tgaggatgtg tgtcagatct gcgctgagaa gttcaaggtg 1860
ggggaccctg aggagtacag cctctttctc ttcgttgacg agacatggca gcagctggca 1920
gaggacactt accctcaaaa aatcaaggcg gagctgcaca gccgaccaca gccccacatc 1980
ttccactttg tctacaaacg catcaagaac gatccttatg gcatcatttt ccagaacggg 2040
gaagaagacc tcaccacctc ctagaagaca ggcgggactt cccagtggtg catccaaagg 2100
ggagctggaa gccttgcctt cccgcttcta catgcttgag cttgaaaagc agtcacctcc 2160
tcggggaccc ctcagtgtag tgactaagcc atccacaggc caactcggcc aagggcaact 2220
ttagccacgc aaggtagctg aggtttgtga aacagtagga ttctcttttg gcaatggaga 2280
attgcatctg atggttcaag tgtcctgaga ttgtttgcta cctacccccca gtcaggttct 2340
aggttggctt acaggtatgt atatgtgcag aagaaacact taagatacaa gttcttttga 2400
attcaacagc agatgcttgc gatgcagtgc gtcaggtgat tctcactcct gtggatggct 2460
tcatccctgc cttccttcct ttctttttcc tttgtgtgtt ttttttttt ttttaaaaaa 2520
gccttcgggt tttaaa 2536

<210> 70
<211> 1415
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1901373CB1

<400> 70
gcgaggacgc gggccgagcc ggaagtggag tgcgctgcgg cgcgagctgg gccggcgggc 60
gtggttcgag agcgcgcaga gtccagactg gcggcagggc ccgaggggcc gacccgcagc 120
gtccctggtc tctccagccc tcactcggaa ccgcactgac aataccctcc cctcccttgg 180
gctggacccc tctctacagc taggagccaa tggcagaaga caaaaccaaa ccgagtgagt 240
tggaccaagg gaagtatgat gctgatgaca acgtgaagat catctgcctg ggagacagcg 300
cagtgggcaa atccaaactc atggagagat ttctcatgga tggctttcag ccacagcagc 360
tgtccacgta cgccctgacc ctgtacaagc acacagccac ggtagatgga aggaccatcc 420
ttgtggactt ttgggacacg gcaggccagg agcggttcca gagcatgcat gcctcctact 480
accacaaggc ccacgcctgc atcatggtgt ttgatgtaca gaggaaagtc acctatagga 540
acctgagcac ctggtataca gagcttcggg agttcaggcc agagatccca tgcatcgtgg 600
tggccaataa aattgatgca gacataaacg tgacccaaaa aagcttcaat tttgccaaga 660
agttctccct gcccctgtat ttcgtctcgg ctgctgatgg taccaatgtt gtgaagctct 720
tcaatgatgc aattcgatta gctgtgtctt acaaacagaa ctcccaggac ttcatggatg 780
agatttttca ggagctcgag aacttcagct tggagcagga agaggaggac gtgccagacc 840
aggaacagag cagcagcatc gagaccccat cagaggaggt ggcctctccc cacagctgag 900
gggctggggc taggggtggg tggagccctt ttaaaatacc cttcccttca acaactctcc 960
agctctgaat ggagaaactc tctaggccat cccctcttct acctcctgca acccacccat 1020
cctattagcc tcccacattc aaggcccgtg atacagggat gaggtcagca ccagcaaact 1080
ctggactggt ggaagaattc cccaccagat ctccttgaag cagaattagg gatcagcatc 1140
attaacacct tccccacccc ctccccgcag gcagacagtg aagagaatca gaaaacatga 1200
ttatgtgtca ctttaataca ggaaatttag gtgtttttttg gtgtttttgt ttttgttttt 1260
gtttctttc caaagctcac ctcggggaca attccttggg cttctcctga ggtaatgatt 1320
taccccccca cccacagctg agtctgtgag gccccatcct ttccctacgt tttctcccat 1380
cttttttcct cttcagtctc ccagtcatct ggttt 1415

<210> 71
<211> 1902
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2367767CB1

<400> 71
gcgaggtctg gctaggctac gggccacgcg ccgccgccgc tgccgccgcc actgtcctct 60
tcggaggcgc gggcccgacg gaaaccatgt ttgtggctcg cagcatcgcg gcggaccaca 120
aggatctcat ccacgatgtc tctttcgact tccacgggcg gcggatggca acctgctcca 180

gcgatcagag cgttaaggtc tgggataaaa gtgaaagtgg tgattggcat tgtactgcta 240
gctggaagac acatagtgga tctgtatggc gtgtgacatg ggcccatcct gaatttgggc 300
aggttttggc ttcctgttct tttgaccgaa cagctgctgt atgggaagaa atagtaggag 360
aatcaaatga taaactgcga ggacagagcc actgggttaa aggacaact ctggtggata 420
gcagaacatc tgttactgat gtgaagtttg ctcccaagca catgggtctt atgttagcaa 480
cctgttccgc agatggtata gtaagaatct atgaggcacc agatgttatg aatctcagcc 540
agtggtcttt gcagcatgag atctcatgta agctaagctg tagttgtatt tcttggaacc 600
cttcaagctc tcgtgctcat tcccccatga tcgccgtagg aagtgatgac agtagcccca 660
acgcaatggc caaggttcag atttttgaat ataatgaaaa caccaggaaa tatgcaaaag 720
ctgaaactct tatgacagtc actgatcctg ttcatgatat tgcattcgct ccaaatttgg 780
gaagatcttt ccatattcta gcaatagcga ccaaagatgt gagaattttt acattaaagc 840
ctgtgaggaa agaactgact tcctctggtg ggccaacaaa gtttgaaatc catatagtgg 900
ctcagttcga taatcataat tctcaggtct ggcgagtgag ttggaatata acaggaacgg 960
tgctagcatc ttcaggagat gatgggtgtg taagattgtg gaaagctaat tatatggaca 1020
attggaagtg tactggtatt ttgaaaggta atgggagccc agtcaatggg agttctcagc 1080
agggaacctc aaatccttcc ctaggttcaa atattccaag tcttcagaat tcattaaatg 1140
gatcttctgc tggcagaaag cacagctgag tacaagctaa ctggagtaac tttgctgttt 1200
tgctgcttgt tgcatgcaca caggaatgga aagcgagctc cttttcccct tccccagcgc 1260
cgtttgacct ctcccaagat acaccagcag cctgcttact actaaacgca atccaaaagg 1320
cctttaaaaa tacagtgtat atttttgta ctagtcagtt tattgacact atttgaaact 1380
tttgaaatat aaacggagag gctttctgtt gagacattgt caccaaaaca attttttgaa 1440
atgttcctga aactaatttg ggtttaaaga ttaaaagggt tgttaccatt cttatctgag 1500
tagttgggag gaggggaata ccactttagt tcatttggaa aatatagaca tatttctttt 1560
gctttcttaa aacagcttaa aatgatgaac ttttataatt ttaatttgaa gattgaataa 1620
atattttta taaagattgt tttgagtgct gatttgttta ctttttgtag atttgcttta 1680
tccatgatat tcagtacaac tctgtcattt ctttgtaata tttaaaaaat attagtaaag 1740
gagtgaatta ataaagtagt aatagtaaaa tgaaaggaac ttgactgtac agtttgtagc 1800
caggttaagc atttggtatt gtttcattta caatttggga ctaagatgga aacacttttt 1860
ttataagttt ttaattcata gtcactaaag agataaatgt tt 1902

<210> 72
<211> 1681
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3090433CB1

<400> 72
gggcagggct ttgctatggc taatgatccc ttggaaggct tccatgaagt aaaccttgct 60
tcacctactt ctccggacct tcttggtgtg tatgaatcag gaactcaaga gcagactacc 120
tcaccaagtg tcatctaccg gccacaccct tcagctttat cctctgtacc tatccaggca 180
aatgcattag atgtttctga acttcctaca caacccgtgt attcatcccc cagacgttta 240
aattgtgcgg aaatatctag tatcagcttt catgttacag acccagcccc ttgctctacc 300
tctggagtca cagctggatt aactaaatta actacaagaa aggacaacta taatgcagag 360
agagagtttt tacagggtgc tactataaca gaggcttgcg atggcagtga tgatattttt 420
gggttgagta ctgatagtct gtctcgttta cgaagcccat ctgttttgga agttagagaa 480
aagggctatg aacgattaaa agaagaactc gcaaaagctc agagggaact gaagttaaaa 540
gatgaagaat gtgagaggct ttcaaaagtg cgagatcaac ttggacagga attggaagaa 600
ctcacagcta gtctatttga ggaagctcat aaaatggtga gagaagcaaa tatcaagcag 660
gcaacagcag aaaaacagct aaaagaagca caaggaaaaa ttgatgtact tcaagctgaa 720
gtagctgcat tgaagacact tgtattgtcc agttctccaa catcacctac gcaggagcct 780
ttgccaggtg gaaagacacc ttttaaaaag gggcatacaa gaaataaaag cacaagcagt 840
gctatgagtg gcagtcatca ggacctcagt gtgatacagc caattgtaaa agactgcaaa 900
gaggctgact tatccttgta taatgaattc cgattgtgga aggatgagcc cacaatggac 960
aggacgtgtc ctttcttaga caaaatctac caggaagata tctttccatg tttaacattc 1020
tcaaaaagtg agttggcttc agctgttctg gaggctgtgg aaaacaatac tctaagcatt 1080
gaaccagtgg gattacaacc tatccggttt gtgaaagctt ctgcagttga atgcggagga 1140
ccaaaaaaat gtgctctcac tggccagagt aagtcctgta aacacagaat taaattaggg 1200
gactcaagca actattatta tatttctcct ttttgcagat acaggatcac ttctgtatgt 1260
aacttttta catacattcg atacattcag cagggactcg tgaaacagca ggatgttgat 1320
cagatgtttt gggaggttat gcagttgaga aaagagatgt cattggcaaa gctgggttat 1380

ttcaaagagg aactctgatg ctctgcgtgg gaccatgcct gaactccccg aataactgaa 1440
aaatggctga atatttttat ggttacttga tatttatttc caaggagtga gcctaagact 1500
tttttcccct tttgcaaatt gctctaagaa gtaccatgat ttcttttaaa ctgatctatg 1560
ctgtgtttgc ttattcttta gttgaacaca ctatgaagaa ttccaggtgt actagtgaat 1620
gtaatttata gttgccaaaa aaaaaacaaa cctgaaataa ataaatgtta gattgaaaaa 1680
a 1681

<210> 73
<211> 1378
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3800591CB1

<400> 73
ggcagatcct atctggcgca tgcgaacgct tctgtgccga ttccttgaag agcaggcgca 60
gactcaaggc tgttgcttcc gcccttactc cccgccgctc gtccctgggc ggggcgaagg 120
ctgggctggg ggaagaggcg tggcggcgct gtgcgcgtgc acaaaagaga gctgaggggc 180
gggggcgctg cggcacagct ggtttgagca actgaactgg aaacaagatg caggacccca 240
acgcagacac tgaatggaat gacatcttac gcaaaaaggg tatcttaccc cccaaggaaa 300
gtctgaaaga attggaagag gaggcagaag aggagcagcg catcctccag cagtcagtgg 360
tgaaaacata tgaagatatg actttggaag agctggagga tcatgaagac gagtttaatg 420
aggaggatga acgtgctatt gaaatgtaca gacggcggag actggctgag tggaaagcaa 480
ctaaactgaa gaataaattc ggagaagttt tggagatctc agggaaggat tatgttcaag 540
aagttaccaa agctggcgag ggcttgtggg tcatcttgca cctttacaaa caaggaattc 600
ccctctgtgc cctgataaat cagcacctca gtggacttgc caggaagttt cctgatgtca 660
aatttatcaa agccatttca acaacctgca tacccaatta tcctgatagg aatctgccca 720
cgatatttgt ttacctggaa ggagatatca aggctcagtt tattggtcct ctggtgtttg 780
gcggcatgaa cctgacaaga gatgagttgg aatggaaact gtctgaatct ggagcaatta 840
tgacagacct ggaggaaaac cctaagaagc cgattgaaga cgtgttgctg tcctcagtgc 900
ggcgctctgt cctcatgaag agggacagcg attccgaggg tgactgaggc tacagcttct 960
atcacatgcc gaactttctt gtgacaaatt gtctggattt tttaaaaaag gaaaaagcaa 1020
gaatgaatcc ttgtggtttt tagttttgta taaattatgt ttcaaatctt tacattttgg 1080
aaataatcat tgctggagat tctgttaaat attttggaac tctttttttt ttaaattata 1140
gtatttcctc taaaaaaaat taaaaccagc catttgtatg gcaaaaaaaa aaaagatact 1200
tcaatattac aattcaggtt tcctaatttt ctaaaaccta tgggaatttt ctaggatgga 1260
cgatcttagg aaggatcact tttggctgtt gtgagaaaca caaaataatt ttattacact 1320
ttaaaaatgt tttgtcataa tttagttaat attaaccttg tttaacttta tagaaaga 1378

<210> 74
<211> 1444
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5308471CB1

<400> 74
gcacgcagtg ccggaggccg cagcgccgga acctcagagg cgggtcgcag cggcgcagag 60
gaggtcagct gcgggagcgt ttccggggac ggtgccgcca tgagattgac cccgcgcgcg 120
ctgtgcagcg ccgcccaggc cgcctggcgg gagaacttcc ccctgtgcgg tcgcgacgtg 180
gcgcgctggt tcccgggcca catggccaag ggcctgaaga agatgcagag cagcctgaag 240
ctggtggact gtatcatcga ggtccacgat gcccggatcc cactttcagg ccgcaaccct 300
ctgtttcagg aaacccttgg gcttaagcct cacttgctgg tcctcaacaa gatggacttg 360
gcggatctta cagagcagca gaaaattatg caacacttag aaggagaagg cctaaaaaat 420
gtcattttta ccaactgtgt aaaggatgaa aatgtcaagc agatcatccc gatggtcact 480
gaactgattg ggagaagcca ccgctaccac cgaaaagaga acctggagta ctgtatcatg 540
gtcattgggg tccccaacgt gggcaagtcc tccctcatca actccctccg gaggcagcac 600
ctcaggaaag ggaaagccac cagggtgggt ggcgagcctg ggatcaccag agctgtgatg 660
tccaaaattc aggtctctga gcggcccctg atgttcctgt tggacactcc tggcgtgctg 720

gctcctcgga ttgaaagtgt ggagacaggc ctgaagctgg ccctgtgtgg aacggtgctg 780
gaccacctgg tcggggagga gaccatggct gactacctgc tgtacaccct caacaaacac 840
cagcgctttg ggtacgtgca gcactacggc ctgggcagtg cctgtgacaa cgtagagcgc 900
gtgctgaaga gtgtggctgt gaagctgggg aagacgcaga aggtgaaggt gctcacgggc 960
acgggtaacg tgaacgttat tcagcctaac tatcctgcgg cagcccgtga cttcctgcag 1020
actttccgcc gtgggctgct gggttccgtg atgctggacc tcgacgtcct gcggggccac 1080
cccccggctg agactttgcc ctgaacttgt ccgggtaggg agggccggag gcatgtggcc 1140
tcccagacct cctgacctgg gtggttgagg ctcaagacag ctcacccggt ccagaagctc 1200
catgctggtc actagggtgc tgtgctctct ggcgccccac agcctggcca gctccaggga 1260
ccccagttgc agggcccaag caggtgggag tggacaccag gcttcccagt ggacgtccct 1320
gagcagctcc gcatgcttgg ttctcccgga gcttcctgct caggcctctt gagaaatgga 1380
tgctgtctca gaaggagtta aagctataac ctgtaacctt taaaatctca aaaaaaaaaa 1440
aaaa 1444

<210> 75
<211> 2067
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5324322CB1

<400> 75
ggcactgtcc accagcactc agagctccat tatgtcccca gctggggttg cagggtaggg 60
gggactgggg tgtcccccag cctcagcaga cggagggcct cagggatgag gctgccagga 120
tagcgccaga gaagcagctc agagcaaggg ctcctgagtg gggcagggc tggggagaag 180
gtcatggggg ggctgcagta ggggtggtca ttgtgcaggc tgagttgaga gaagtgggtg 240
gccatgttct cctcagacag aaactgcttg cgcagaggct ccctggggag agatggcaga 300
gaggcaggct gggatactga cacaggaggc agcctgttgg ggaccagagg tgacagagat 360
cttgttggga gtccctccct gcccccaaac tcactgctcc tcctccaggc gccgcttggt 420
gctcatgggc acagctcctc ggagagggga gctggcgtcc aggccccaag tcacccccaa 480
ggcggcccgc gggaggcgct gggcccctcc ctggggggcct cgctgcaagg gctgctgcag 540
gatcattggg ttttggggtc ctgcgggtgg gatctgggcg acaggggagg agtctctgag 600
ggcgtggcca agagaggatg ggcgtggctt taggcgggca cagccgcgag gttctgcgcg 660
ggcgcggaag acgggcggcg cgtggcggaa ggcaggcttg ctcctcgggg tgggggaggg 720
tatccggctt aagggggctg cggtggacac cacttcttaa tgtcgggggt cttcgcggcg 780
ctcacctcgg ctcctagggt tcgggacggt acgcaccagc caccttcgcg ccgaaggcgg 840
tagggcgcca cggagaggaa ccgctctagg cacgtaaggc ctcgtgaggt tgcgtcgcgc 900
gcggagcact ctgggacttg tagttctgga gatggagcga gctgtgccgc tcgcggtgcc 960
tctgggtcag acagaggtgt tccaggcctt gcagcggctc catatgacca tcttctccca 1020
gagcgtctca ccatgtggga agtttctggc ggctggcaac aattacgggc agattgccat 1080
cttcagcttg tcctctgctt tgagctcaga agccaaagag gaaagtaaga agccggtggt 1140
gactttccaa gcccatgatg ggcccgtcta tagcatggtt tccaccgatc gacatctgct 1200
tagtgctggg gatggggagg tgaaggcctg gctttgggcg gagatgctca agaagggctg 1260
taaggagctg tggcgtcgtc agcctccata caggaccagc ctggaagtgc ctgagatcaa 1320
cgctttgctg ctggtcccca aggagaattc cctcatcctg gctgggggag actgtcagtt 1380
gcacactatg gaccttgaaa ctgggacttt cacgagggtc ctccggggcc acacagacta 1440
catccactgc ctggcactgc gggaaaggag cccagaggtg ctgtcaggtg gcgaggatgg 1500
agctgttcga ctttgggacc tgcgcacagc caaggaggtc cagacgatcg aggtctataa 1560
gcacgaggag tgctcgaggc cccacaatgg gcgctggatt ggatgtttgg caactgattc 1620
cgactggatg gtctgtggag ggggcccagc cctcaccctc tggcacctcc gatcctccac 1680
acccaccacc atcttcccca tccgggcgcc acagaagcac gtcaccttct accaggacct 1740
gattctgtca gctggccagg gccgctgcgt caaccagtgg cagctgagcg gggagctgaa 1800
ggcccaggtg cctggctcct ccccagggct gctcagcctc agcctcaacc agcagcctgc 1860
cgcgcctgag tgcaaggtcc tgacagctgc aggcaacagc tgccgggtgg atgtcttcac 1920
caacctgggt taccgagcct tctccctgtc cttctgatct ctgacgacac cccagccag 1980
ctcagggttt tagagtgttt ttcattttct tttttttttt ttttttacaa taaagtttca 2040
ggctttttta ccaaaaaaaa aaaaaaa 2067

<210> 76
<211> 2085
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 067184CB1

<400> 76
gtgttgcgcg actggccttg agggagagct ggggcctgct cccggagaga tacggctatg 60
tcgatcgaaa tcgaatcttc ggatgtgatc cgccttatta tgcagtactt gaaggagaac 120
agtttacatc gggcgttagc caccttgcag gaggagacta ctgtgtctct gaatactgtg 180
gacagcattg agagttttgt ggctgacatt aacagtggcc attgggatac tgtgttgcag 240
gctatacagt ctctgaaatt gccagacaaa accctcattg acctctatga acaggttgtt 300
ctggaattga tagagctccg tgaattgggt gctgccaggt cacttttgag acagactgat 360
cccatgatca tgttaaaaca aacacagcca gagcgatata ttcatctgga gaaccttttg 420
gccaggtctt actttgatcc tcgtgaggca tacccagatg gaagtagcaa agaaaagaga 480
agagcagcaa ttgcccaggc cttagctggc gaagtcagtg tggtgcctcc atctcgtctc 540
atggcattgc tgggacaggc actgaagtgg cagcagcatc agggattgct tcctcctggt 600
atgaccatag atttgtttcg aggcaaggca gctgtcaaag atgtggaaga agaaaagttt 660
cctacacaac tgagcaggca tattaagttt ggtcagaaat cacatgtgga gtgtgctcga 720
ttttctccag atggtcagta tttggtcact gggtctgttg atggattcat tgaagtatgg 780
aactttacta ctggaaaaat cagaaaggat cttaagtacc aggcccaaga taactttatg 840
atgatggatg atgctgtcct ctgcatgtgt ttcagcagag atacagaaat gttagcaact 900
ggggcccaag atggaaaaat caaggtgtgg aagattcaga gtggacaatg tttaaggaga 960
tttgagaggg cacacagtaa gggtgtcacc tgtctaagct tttctaagga tagcagtcag 1020
atccttagtg cttcttttga ccagacaatt agaattcatg gtttaaaatc tgggaaaacc 1080
ctgaaggaat ttcgtggcca ttcctccttt gttaacgaag caacatttac acaagatgga 1140
cattacatta ttagtgcatc ctctgatggc actgtaaaga tctggaatat gaagaccaca 1200
gaatgttcaa atacctttaa atccctgggc agcaccgcag ggacagatat taccgtcaac 1260
agtgtgattc tacttcctaa aaaccctgag cactttgtgg tgtgcaacag atcaaacacg 1320
gtggtcatca tgaacatgca ggggcagatt gtcagaagct tcagttctgg taaaagagaa 1380
ggtggggact ttgtttgctg tgccctctct ccccgtggtg aatggatcta ctgtgtaggg 1440
gaggactttg tgctctactg tttcagtaca gtcactggca aactggagag aactttgaca 1500
gtgcacgaga aggatgtgat tggtattgca catcaccctc atcagaacct gattgctacc 1560
tacagtgaag atggactcct aaagctctgg aaaccataat tcaacttttc tttttaaatc 1620
agctcgaaag catgtactta aatgaagcat attcatgtaa tgtgcttttt tttttttttt 1680
gccagctttt ctaagcaaat agattgtctg aattagtcac agaataattt tgtgaaaatt 1740
catgtttaag tagcaactac cctttctttt tttatatatt tttaaggtat tagtttatct 1800
tcttctaact ggtgcagtca cttaatgttt tcattaatct tcgacctgga gagtgaaata 1860
ctgatatttc tagaaaaaaa ttctactcct ctgattattt gaaatgctga ggaaaatgtc 1920
cctcccatag taaaacttgt aaataaggaa ctatatcata ttcagtagct gtgttctgtt 1980
ccatcttttt tttttttttt gagatggagt tttgcttgtt gcccaggctg gagtgcagcg 2040
gcacgatctt ggttcactgc aacctccgcc tcccaggttc aagcg 2085

<210> 77
<211> 2061
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 722896CB1

<400> 77
cgagccgcgg gacccgggcc gtaccgggga ggggccgctc cgggccgcag cgcgagggca 60
gcgaggggcg gcggggacct ggcaccgggc ggggccggcg gcagcgacca tgatcgcttt 120
gttcaacaag ctgctggact ggttcaaggc cctattctgg aaggaggaga tggagctcac 180
gctggtcggg cttcagtact cgggcaagac caccttcgtc aacgtgatcg cgtcaggaca 240
gttcaacgag gacatgatcc ccaccgtggg tttcaacatg cgcaaaatca ccaaagggaa 300
tgtgactatc aagctctggg acattggggg acagccgcgt ttccgcagca tgtgggagcg 360
ctactgccga ggagtgagcg ccatcgtgta catggtggat gctgctgacc aggagaagat 420
tgaggcctct aagaacgagc tccacaacct actggacaaa cctcagctgc agggcatccc 480
ggtcttagtc ctgggtaaca agcgagacct tccgggagca ttggatgaga aggagctgat 540
tgagaaaatg aatctgtctg ccatccagga ccgagagatc tgctgctact ccatctcttg 600

caaagaaaag gacaacattg acatcaccct acagtggctt attcaacact cgaagtcacg 660
gagaagctga gactccagcc cttctccctc agaccaggga ccgtcatcat ctaaacctga 720
agccgagctc cccgcccacc cctgtcgtcc ccctaagccc acccctcctc acccagtgtg 780
aggagggccc tctggggacc ccagagtcct gttctgctga ggtttgaact cctgttttta 840
ttgtaaaata aattgccccc cattctggtc ccctaacttc tcacccttcc ccgctgcctt 900
tgtcccatca cccagccctg cctccctccc agcagccctg ggccacagcc cccgcccctg 960
gcttttcccc ggcccggtct tgtacctccc ttttcaacac tctctgttat tgtcctgtgt 1020
gtacagtata tatatgtata tatattttaa ttttttaatt taagcaaaga ctaaaatcaa 1080
ccatttgatg ctgcaggggc ctttcaggat ctgggagggg gcagtctgga gagaaggagg 1140
gagacgcagg tggacttggg gcaagttcag atcagaagag gtgcaggctg gcacctgcgg 1200
caggtaccag cctgggcact ggtggccgcc tccctgtccc gtgtgtttcc accgcccaat 1260
ctggcttgtc ctggcagtgc ttgaatgcca caggctggca ggggcctctg ggggcccctc 1320
ccctcgaccc ccagcctggg tagagccacc aggtacgacg accaggtacc agaaaccacc 1380
aggcacacgg ggcagaaagc cagcgtccat gccccagcag ccccctcctg cctgttcctg 1440
gctcccagct cccgcccctc cccagggccc ccacctccac ggcccacttc attttctgtt 1500
ctcattttgc agagttgcac aaggagagaa ctcagcatgg ggggttggtt ctttgggttc 1560
tgtttgttta tttgtttaat ttaatgattt gtaaagtgat gttcctcttc ctttttttaca 1620
cttttcagct catatttaac ctctgtttgg aaaatgattc ttgtaactgt acattttttt 1680
gcttcctaat aacaatgaca acaaaaaaaa taaatgacca gttttgtgtt gggggggggtg 1740
tatggtgctg gttacttttc cgcagttggc atgggttgcc ctacaggccc acagggccac 1800
cagcacaccc ccgcacgctg ggcaccaaca gagccacgga gcgcgagcac atgcccgccc 1860
ggggagcaca atggcgctgc acaaaacggc ctcccacacg tgcgtccagg ctcttgcgcc 1920
acctccttct cattctcttt tcagactttc atgtagtccc agctttgagc cagcagctgc 1980
cacttggggc tgcagcgctc tgttgaggga actgcccagg gctgggtaga ggcagcaagg 2040
ggacagggct gggtgctgtt t 2061

<210> 78
<211> 981
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1571739CB1

<400> 78
gtagttccag aaaataggac tgaccaagaa gcagaaaagc aagatgaatg atgtgaagct 60
tgctgtcttg ggtggtgaag gaacaggcaa atctgccctt acagtgaggt ttcttactaa 120
gcgattcatt ggagaatatg cttctaattt tgaatctatc tataagaagc acttgtgttt 180
ggaaaggaaa caactaaatc tagaaatata tgacccttgt tctcaaacac agaaagcaaa 240
attctccctc acaagtgagc ttcactgggc agatgggttt gttattgtgt atgacatcag 300
tgataggtct tcatttgctt ttgcaaaagc gctgatctac agaatccggg agccacaaac 360
tagtcattgt aaaagagctg tggaatcagc agtgtttttg gttggcaaca aacgagatct 420
ttgtcatgtg cgagaggttg gctgggaaga agggcaaaag ctggcactgg aaaaccgatg 480
ccaattctgt gaactgtctg cagcagagca gtctctggag gtggaaatga tgtttatcag 540
aattatcaag gacatcctga taaacttcaa actcaaagaa aagagacgtc ccagtggatc 600
taaatcaatg gccaaattga tcaataatgt atttggaaag agaaggaaat ctgtttagta 660
gacaggtaat cctgggagat ttcctatatc agagagtttc aaacattcac atgataatta 720
aactaacctt tgtatgcaat tttttttttgg taaaaagaat tctcttggag atatgaaatg 780
attgagtatg aaccacagct gtgtttcaa atatgtagtt tgcctttttg gttgttgtac 840
cctgctcact ctccttcaca cagaaccttt catttattgt acaacatcac actcacccta 900
acctactggc ggacagcgat cccagtttgc cttgccaaat aaactctgtt tatgtgaatt 960
tattaaacga caaaaaaaaa a 981

<210> 79
<211> 1375
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1739479CB1

<400> 79
aattctgtgc ttctccctct tgggccttca cacacttatg cttatgtaaa taattattaa 60
tcatattttc atgatggtgg agattcttat tcaccactac atcccatccc cagggcctgc 120
cacagaatag gcattcagta aatatttttt gaatgattga ctgagaaatc acacctctgt 180
ttcttttaaa cacatcctga tagctccata agtttcatca gggtcagtgg tttccattgt 240
cctgactgct ggccacagtg acctgttctg tgctttattt gacaagaccc ctgaatggtt 300
ggacagtgat tcctgccaaa agtgtgatca gcctttcttc tggaacttca agcaaatgtg 360
ggacagtaag aaaattggtc taagacagca ccactgccgc aagtgtggga aggccgtctg 420
tggcaagtgc agctccaagc gctcctccat ccccctgatg ggcttcgagt ttgaagtgag 480
ggtctgtgac agctgccacg aggccatcac agatgaagaa cgtgcaccca cagccacctt 540
ccatgacagt aaacataaca ttgtgcatgt gcatttcgat gcaaccagag gatggttact 600
gacttctgga actgacaagg ttattaagtt gtgggatatg accccagtcg tgtcttgatg 660
actctcccag gaatcagaaa gatagtattt actaaagaaa cggttgtttt aacccaaatc 720
attaccagag tggtaaagca gacatgtgag aagtaagaaa gaaactaaag accctgaatg 780
aatttgcaga ttacccatgt gcacagtggg gacctggcca gtgagcactc gcaaggggac 840
tcttccaact tgttcataca atataaaaga agctattttt ttaacaaatg gtttatacag 900
tctggctgtg ctgcattgtt ttgagtgtac cgaaaaatct gtgtggggtg tttaattttt 960
atacttttca acaccccatt ttatttgttg ctttgtcaga gaaataaggg aggtatctac 1020
tcagagtatt ttggtcatta tactttctgt gtttacttca acatgtgtca cgtggccagc 1080
ggctttttct tctcttccct ctgcacctac ctgcaccttc tctgcctttc ctggagggga 1140
tgtatttatg ttatttattc ccagtgtttc tgctttcatg tcctcctcag tggagagatt 1200
tggaaactca tcatgtggat tcaccagcca gctgctggaa ttgcctgaag agcgatttgt 1260
ttgtaatgtc tgcctcattc acgttcttat gaagtagaaa agactgtgtt tctgcctcag 1320
ttgcctctgt ctttcccaca ttaaaaaaaa aaatgctgtg agaaaaaaaa aaaaa 1375

<210> 80
<211> 2833
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1999147CB1

<400> 80
cggttgggac gcacacactc tgcgtcatgg agggctgagg ccgatgatga attccggagt 60
gcctgtcagg cttgctgtgt cactcggccc gctcggcgcg cccccttccca gccgcccttc 120
cgtaccggct ctcgggctct tccggtctcc ggccgcccct tacctgcagg ctcttctccc 180
gccgcggccc ggcgctctcc gagtcgcccc tgcggactgg tctcgcacag tgcctgggca 240
ccgggcgcca gacagacact ggccatgacg agcggcgcaa ccaggtaccg gctgagctgc 300
tcgctccggg gccacgagct ggacgtacgg ggcctggtgt gctgcgccta tccgccggga 360
gcctttgtgt ccgtgtcccg agaccgcacc acccgcctct gggccccaga cagtccaaac 420
aggagcttta cagaaatgca ctgtatgagt ggccattcca attttgtatc ttgtgtatgc 480
atcatacccct caagtgacat ctaccctcat ggcctaattg ccaccggtgg aaatgaccac 540
aatatatgca ttttctcact ggacagtcca atgccacttt atattctaaa aggccacaaa 600
aatactgttt gtagtctatc atctggaaaa tttgggacat tacttagtgg ttcatgggac 660
accactgcta aagtctggct gaatgacaag tgcatgatga ccttgcaggg tcatacagct 720
gcagtgtggg cggtaaagat cttacctgaa cagggcttaa tgttgactgg atcagcagac 780
aagactgtta aactgtggaa ggctggaaga tgtgagagga ctttttcagg gcatgaagac 840
tgtgtaagag gtttggcaat tttgagtgaa acagaatttc tttcctgtgc aaatgatgct 900
agtattagaa ggtggcaaat cactggcgag tgtcttgaag tatattatgg acatacaaat 960
tatatttata gcatatccgt ttttccaaat tgtagagact ttgtgacaac agcagaggac 1020
agatctctga gaatctggaa acatggggaa tgtgctcaaa ctatccgact tccagctcag 1080
tctatatggt gctgctgtgt gctcgacaat ggtgacattg tggttggtgc gagtgatggc 1140
attattagag tgtttacaga atcagaagat cgaacagcaa gtgctgaaga aatcaaggct 1200
tttgaaaaag aactgtctca cgcaaccatt gattctaaaa ctggcgattt aggggacatc 1260
aatgctgagc agcttcctgg gagggaacat cttaatgaac ctggtactag agaaggacag 1320
actcgtctaa tcagagatgg ggagaaagtc gaagcctatc agtggagtgt tagtgaaggg 1380
aggtggataa aaattggtga tgttgttggc tcatctggtg ctaatcagca aacatctgga 1440
aaagttttat atgaagggaa agaatttgat tatgttttct caattgatgt caatgaaggt 1500
ggaccatcat ataaattgcc atataatacc agtgatgacc cttggttaac tgcatacaac 1560
ttcttacaga agaatgattt gaatcctatg tttctggatc aagtagctaa atttattatt 1620
gataacacaa aaggtcaaat gttgggactt gggaatccca gcttttcaga tccatttaca 1680

ggtggtggtc ggtatgttcc gggctcttcg ggatcttcta acacactacc cacagcagat 1740
ccttttacag gtgctggtcg ttatgtacca ggttctgcaa gtatgggaac taccatggcc 1800
ggagttgatc catttacagg gaatagtgcc taccgatcag ctgcatctaa aacaatgaat 1860
atttatttcc ctaaaaaaga ggctgtcaca tttgaccaag caaaccctac acaaatatta 1920
ggtaaactga aggaacttaa tggaactgca cctgaagaga agaagttaac tgaggatgac 1980
ttgatacttc ttgagaagat actgtctcta atatgtaata gttcttcaga aaaacccaca 2040
gtccagcaac ttcagatttt gtggaaagct attaactgtc ctgaagatat tgtctttcct 2100
gcacttgaca ttcttcggtt gtcaattaaa caccccagtg tgaatgagaa cttctgcaat 2160
gaaaaggaag gggctcagtt cagcagtcat cttatcaatc ttctgaaccc taaaggaaag 2220
ccagcaaacc agctgcttgc tctcaggact ttttgcaatt gttttgttgg ccaggcagga 2280
caaaaactca tgatgtccca gagggaatca ctgatgtccc atgcaataga actgaaatca 2340
gggagcaata agaacattca cattgctctg gctacattgg ccctgaacta ttctgtttgt 2400
tttcataaag accataacat tgaagggaaa gcccaatgtt tgtcactaat tagcacaatc 2460
ttggaagtag tacaagacct agaagccact tttagacttc ttgtggctct tggaacactt 2520
atcagtgatg attcaaatgc tgtacaatta gccaagtctt taggtgttga ttctcaaata 2580
aaaaagtatt cctcagtatc agaaccagct aaagtaagtg aatgctgtag atttatccta 2640
aatttgctgt agcagtgggg aagagggacg gatattttta attgattagt gttttttcc 2700
tcacatttga catgactgat aacagataat taaaaaaaga gaatacggtg gattaagtaa 2760
aattttacat cttgtaaagt ggtggggagg ggaaacagaa ataaaatttt tgcactgctg 2820
aaaaaaaaaa aaa 2833

<210> 81
<211> 1752
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2182085CB1

<400> 81
gcaggcagcc atcttgcctg gagcctgaga aagggaggag agacagaagg aaccggcgac 60
agtggtctca gggccgctcc ggggggcctc aagaaccgga ggcagccccg gaggctgccg 120
cgggcggaca cgccagagga ggaggccggg gaatggccgc ggtgtggcag caagtcttag 180
cagtggacgc gaggtacaac gcgtaccgca caccaacgtt tccacagttt cggacgcagt 240
atatccgccg gcgcagcagc tgctgcggga gaatgccaag gctgggcacc ccccagcgct 300
gcgtcggcag tacctgaggc ttcgggggca gctgctgggc cagcgctacg ggcccctctc 360
cgagccaggc agtgctcgtg cctatagcaa cagcatcgtc cgcagtagcc gcactactct 420
tgaccgcatg gaggactttg aggatgatcc tcgggccctg ggggcccgtg ggcaccgtcg 480
ttctgtcagc agaggctcct accagctgca ggcgcagatg aaccgtgccg tctatgagga 540
caggccccct ggcagcgtgg tgcccacgtc agcagcagag gcaagtcggg ccatggccgg 600
ggacacgtca ctgagcgaga actatgcctt tgcgggcatg tatcatgttt ttgaccagca 660
cgtggatgag gcagtcccaa gggtgcgctt cgccaatgat gaccgacacc gcctggcctg 720
ctgctcactc gacggcagca tctccctgtg ccagctggtg cctgccccac ccacagtgct 780
tcgcgtgcta cggggccaca cccgtggtgt ctccgacttc gcctggtccc tctccaatga 840
catcctcgtg tccacctcac tggatgccac catgcgcatc tgggcctctg aggatggtcg 900
ctgcatccga gagatccctg accccgatag cgctgaactg ctctgctgca ccttccagcc 960
tgtcaacaac aacctcactg tggtggggaa cgccaagcac aacgtgcatg tcatgaacat 1020
ctccacaggc aagaaagtga agggggctc cagcaagctg acaggccgtg tccttgctct 1080
gtcctttgat gcccctggcc ggctgctctg ggcgggtgat gaccgtggca gtgtcttctc 1140
tttcctcttt gatatggcca cagggaagct gaccaaagcc aagcgtttgg tggtgcatga 1200
ggggagccct gtgaccagca tctcagcccg gtcctgggtc agccgcgagg cccgggatcc 1260
ctcactgctc atcaatgctt gcctcaacaa gttgctgctc tacagggtgg tagacaacga 1320
ggggaccctg cagctgaaga gaagcttccc catcgagcag agctcacatc ctgtgcgcag 1380
catcttctgt cccctcatgt ccttccgcca gggggcctgc gtggtgacgg gcagtgagga 1440
catgtgcgtg cacttctttg atgtggagcg ggcggccaag gctgctgtca acaagctgca 1500
gggccacagt gcacctgtgc ttgatgtcag cttcaactgc gacgagagcc tactggcctc 1560
cagtgacgcc agcggcatgg tcatcgtctg gaggcgggag cagaagtagg gtcctgtcgg 1620
ccctgctgct gtcctccatc ccacccctct tactccagcc tcgtgttgta aataaagttt 1680
cggtggtcat gctgagggcc ggctcccagc tctgccgggg acggacaggg cagagggcag 1740
cgggcagctg ca 1752

<210> 82

<211> 1854
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2216640CB1

<400> 82
cccacgcgtc cgcgcaggat ggcggcagca gtggcggacg aggcggtggc gcgcgatgtg 60
cagcggttgc tagtgcagtt ccaggatgag ggcgggcagc tgctgggttc cccgttcgac 120
gtgccccgtgg acatcacccc ggacaggctg cagctcgtgt gcaacgcgct actggcccag 180
gaggatcccc tgccactggc tttctttgtc cacgatgctg agatcgtctc ctcactgggg 240
aagacgttgg agtcccaggc agtggagaca gagaaggtcc tagacatcat ctaccagcca 300
caggctatct tcagagtccg ggctgtgact cgctgcacca gctccttgga gggtcacagt 360
gaggcagtca tttctgtggc cttcagccct acgggaaagt acctggccag tggctctgga 420
gacaccaccg tgcgcttctg ggatctcagc acagagacac cacatttcac atgcaaggga 480
cacagacact gggtccttag tatatcctgg tctccagatg gcaagaagct ggcctcaggc 540
tgcaagaatg gccagattct cctctgggac ccaagcacag ggaagcaggt gggcaggacc 600
ctcgctggcc acagcaagtg gatcacaggc ctgagctggg agcccctcca tgcgaacccct 660
gagtgccgct atgtggccag cagctccaag gatggcagtg tgcggatctg ggacacaact 720
gcaggccgct gtgagcgcat cctcaccggg cacacccagt cggtcacctg tctccggtgg 780
ggaggggacg ggcttctcta ctctgcctcc caggaccgca ccatcaaagt ctggagagct 840
catgacggtg tgctgtgccg gactctgcaa ggccacggcc actgggtgaa caccatggcc 900
ctcagcactg actatgccct gcgcactggg gcctttgaac ctgctgaggc ctcagttaat 960
ccccaagacc tccaaggatc cttgcaggag ttgaaggaga gggctctgag ccgatacaac 1020
ctcgtgcggg gccagggtcc agagaggctg gtgtctggct ccgacgactt caccttattc 1080
ctgtggtccc cagcagagga caaaaagcct ctcactcgga tgacaggaca ccaagctctc 1140
atcaaccagg tgctcttctc tcctgactcc cgcatcgtgg ctagtgcctc ctttgacaag 1200
tccatcaagc tgtgggatgg caggacgggc aagtacctgg cttccctacg cggccacgtg 1260
gctgccgtgt accagattgc gtggtcagct gacagtcggc tcctggtcag cggcagcagt 1320
gacagcacac tgaaggtgtg ggatgtgaag gcccagaagc tggccatgga cctgcccggc 1380
cacgcggatg aggtatatgc tgttgactgg agtccagatg gccagagagt ggcaagtggt 1440
gggaaggaca aatgcctccg gatatggagg agatgagacg gcccgaagtt ctctctgacc 1500
cccacctcga ctcggcctct gccagctgcc ttccctgcca gagaacaaag gctgagatgg 1560
cagtgcacac accctcccca ccagtgggga cctgagaatg cgtgtggcct gctgtcctcg 1620
atagaccgga atggggtttt cccacagatc cccgcctgtg gcacacccca gagccagaaa 1680
tcgaaggtca caggaagttg tcactgaact tggcccgtgt ctgctactct gtaccttgct 1740
ggtacagaca ggggtggtgg gcagccaggc tctatgagtg ggcccctagt gtcagctctg 1800
tacagggtca gatcccaggt tctatgacca aataagtaac ttaaaaaaaa aaaa 1854

<210> 83
<211> 862
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2417361CB1

<400> 83
ggcgtggctt caacagactt tcttttgcct gtctttgtcc cagagcctct tccctggccc 60
tgctgagacc actgctctaa gaagagacca ccagactgag agaggactcc cagctgccct 120
cagagcggag gccgagtgct gcacagccac agctgctctg aagcccttcc atgaatcccc 180
ggaagaaggt ggacctgaaa ctcattatcg tcggagccat tggtgtggga aagacctccc 240
tccttcacca atatgtgcac aagacgtttt atgaggaata ccagaccaca ctgggggcca 300
gcatcctctc caagattatc atattgggtg acacaacttt gaagttacag atctgggaca 360
cgggcggtca ggagcggttc cgctccatgg tgtccacgtt ctacaagggc tccgatggct 420
gcatcctagc ttttgatgtc accgacctgg agtcttttga agccctggat atctggcggg 480
gtgatgtcct ggccaagatt gtccccatgg agcagtccta ccccatggtg ttgttgggga 540
acaagatcga tctggcagac cggaaggtac cccaggaagt agctcaaggc tggtgtagag 600
agaaagatat tccttacttt gaagtcagtg ccaagaatga catcaatgtg gtgcaagcgt 660
ttgagatgct ggccagtagg gctctgtcga ggtaccagag catcttagaa aatcacctca 720

cagaatccat caagctctcg ccagaccagt caaggagcag atgctgctga cctccagacg 780
cctgctctgg aagcccagaa acagagcctg ccccgagcct ggtcacccca ggcttgagaa 840
caggtgacca tccccctcca gc 862

<210> 84
<211> 1406
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2454384CB1

<400> 84
ctagagcctg gggtctcggc aacttccggc ggcgggagct gcagagcgca aggcccgccc 60
actgcgcgtg cgcttcggcc cggctcctcc tgcgcccccg gcccctgcga ctgggacttg 120
gtacggccgg gcggttggcg tcctctgcgg ctcctgccag gggcgggctt ttcaaatctt 180
ccctttgaag gagtggcgac ggcccggaca gttcgcgttg gagatggagg ggccgagcct 240
gaggggtcct gcgctccgcc tggcggggct tcccacccag caggactgca acattcaaga 300
aaaaatagac ttagaaattc gaatgcgaga aggaatatgg aaactccttt ctctgagcac 360
tcagaaagat caagttttac atgcagttaa gaatctcatg gtgtgcaatg ctcgactaat 420
ggcctataca tcggagctac agaaattaga agaacagatt gcaaatcaga ctggaagatg 480
tgatgtgaaa tttgaaagta aagaacgaac agcatgtaaa ggaaagattg ccatatcaga 540
tattcgaata ccactaatgt ggaaagactc tgatcacttc agcaataaag aacgatcacg 600
acgctatgcc attttttgtt tattcaaaat gggagctaat gtgtttgata ctgatgtggt 660
gaatgtggat aaaacaatca cagatatatg ttttgaaaat gtaaccatat tgtaagtatt 720
ttttaatctt cagagaataa aaataattta aaattcttct tttttaaaag aaagttctta 780
ttattggttc tttggattca ttttatgttt aaatgtttaa gtgatcttta aatgtttaat 840
atgattttaa aaattatttt gttcagaaga agtccatttc tctatctgca gttttctgat 900
gtgaaataaa aatggaaatc ttgtaattac tattagcagt aaatatttga cttattagat 960
atgacccatt tttaaattgt taataaatat agttcagtta ttaacaaagc tatgcataca 1020
acagaatatc ctgtaatgtt atttgatata gagagaattt aagcataaaa caggattttt 1080
atctcatgta ggatatttgg ttgcagaaat actaaaatag tatagcgact ttatttacaa 1140
gatagtcctg aagtacatgc tatataggaa gagcactttg aaattttggg gtgttctttt 1200
tcttatggtg cacttctttc atgtacttca aagcaataaa aaaaaatggg tgatctcagg 1260
gctgttttta ttgtccctgc tcttttacag gctcatttta ttgtggtcat aatacagaac 1320
aagaaggaac tccttgggta gccatagaaa tcatttttaa cttacatagt ttttcctgcc 1380
ctccttcaaa ggttctatgt gcctaa 1406

<210> 85
<211> 1184
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2610262CB1

<400> 85
gcggttttgg tgcctgaagc agggagcgcg gagtcgttcc cgagagaggc ggccaggcta 60
tgctcgccgg tttccggcgt tccgctccgg ccagccagag tctctgtctc aacctgtgtc 120
cgtgctccag cagtctcctc agcccggccc cgcggcgcgg ttggcggcgg cgcccccaggc 180
gcgccccctc ctccgatggc ggcggagatc cagcccaagc ctctgacccg caagccgatc 240
ctgctgcagc ggatggaggg gtcccaggag gtggtgaata tggccgtgat cgtgcccaaa 300
gaggagggcg tcatcagcgt ctccgaggac aggacagttc gtgtttggtt aaagagagac 360
agtggacagt attggccaag cgtataccat gcaatgcctt ctccatgttc atgcatgtct 420
tttaacccgg aaacaagaag actgtccata ggtctagaca atggtacaat ctcagagttt 480
atattgtcag aagattataa caagatgact cctgtgaaaa actatcaagc gcatcagagc 540
agagtgacga tgatcctgtt tgtcctggag ctggagtggg tgctgagcac aggacaggac 600
aagcaatttg cctggcactg ctctgagagt gggcagcgcc tgggaggtta tcggaccagt 660
gctgtggcct caggcctgca atttgatgtt gaaacccggc atgtgtttat cggtgaccac 720
tcaggccaag taacaatcct caaactggag caagaaaact gcaccctggt cacaacattc 780
agaggacaca caggtggggt gaccgctctc tgttgggacc cagtccagcg ggtgttgttc 840

tcaggcagtt cagatcactc tgtcatcatg tgggacatcg gtgggagaaa aggaacagcc 900
atcgagctcc aaggacacaa cgacagagtc caggccctct cctatgcaca gcacacgcga 960
caattgatct cctgtggcgg tgatggtggg attgtcgtct ggaacatgga cgtggagagg 1020
caggagcctc tgtggagctg cttcgtggtt atgataagtg ctgtgtgatg ctcaccttgg 1080
gaggtctgcg acatatattg aagtcatctc taacctgaag tactgacaga ctttctggaa 1140
gaaaaggctt gtaggaggaa acttcagaat tctattaaat ggtg 1184

<210> 86
<211> 2965
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2700075CB1

<400> 86
ggcaccactg tgaaggtctg ggacgcagcc aagcagcagc ccctgacaga gctggcagcc 60
catggggacc tggtgcagag cgccgtctgg agccgagatg gagccctggt gggcacggcg 120
tgcaaggaca agcagctgcg gatctttgac cccagaacaa agccgcgggc ctctcagagc 180
acgcaggccc atgagaacag cagggatagc cggctggcat ggatgggcac ctgggagcac 240
cttgtgtcta ctggattcaa ccagatgcgt gagcgcgaag tgaagctgtg ggacacgcgg 300
ttcttctcca gcgccctggc ctccctcacc ttggacacct cgcttgggtg tctcgtgcct 360
ctgctggacc ctgactctgg gctcctggtc ctggcaggaa agggcgagag gcagctgtac 420
tgttacgagg tggtcccgca gcagccggcg ctgagcccag tgacccagtg tgtcctggag 480
agcgtgctgc gtggggctgc ccttgtgccc cggcaggcgc tggccgtcat gagctgcgag 540
gtactccgcg tcctacagct gagcgacaca gccatcgtgc ccatcggcta ccatgtgccc 600
cgcaaggctg tggagttcca cgaggacctg ttcccggaca ctgccggctg tgtgcctgcc 660
accgaccccc atagctggtg ggctggggac aaccagcagg tgcagaaggt cagcctcaac 720
cccgcctgcc ggccccaccc gagcttcact tcctgtctgg tgccccctgc ggagcccctc 780
cctgacacag cccagcctgc ggtgatggag acacccgtgg gtgatgcaga cgcaagcgag 840
ggtttctctt cccctcccag ttcgctgacc tcgccctcca cgccctccag cctgggggcc 900
tcactctcca gcaccagtgg catcgggacc agccccagtt tgaggtcgct gcagagcctg 960
ctgggcccca gttccaagtt ccgccatgct cagggcactg tcctgcaccg agacagccac 1020
atcaccaacc tcaaggggct caacctcacc acacctggtg agagtgacgg cttctgtgcc 1080
aacaagctgc gtgtggccgt gccgctgctc agcagcgggg gacaggtggc tgtgcttgag 1140
ctacggaagc ctggccgcct gcccgacacg gcactgccca cgctgcagaa tggggcagct 1200
gtgactgatc tggcctggga cccctttgac ccccatcgcc tcgctgtggc tggtgaggac 1260
gccaggatcc gactgtggcg ggtacccgca gagggcctgg aagaggtgct caccacgcca 1320
gagactgtgc tcacaggcca cacggagaag atctgctccc tgcgcttcca cccactggca 1380
gccaatgtgc tggcctcgtc ctcctatgac ctcactgttc gcatctggga ccttcaggct 1440
ggagctgatc ggctgaagct gcagggccac caagaccaga tcttcagcct ggcctggagt 1500
cctgatgggc agcagctggc cactgtctgc aaggatgggc gtgtgcgggt ctacaggccc 1560
cggagtggcc ctgagcccct gcaggaaggc ccagggccca agggaggacg cggagctcgc 1620
attgtctggg tatgtgatgg tcgctgtctg ctggtgtctg gctttgacag ccaaagtgag 1680
cgccagctgc tcctatatga agctgaggcc ctggccggcg gacccttggc agtgttgggc 1740
ctggacgtgg ctccctcaac cctgctgccc agctacgacc cagacactgg cctggtgctc 1800
ctgaccggca agggcgacac ccgtgtattc ctgtacgagc tgctccccga gtcccctttc 1860
ttcctggagt gcaacagctt cacgtcgcct gacccccaca agggcctcgt cctcctgcct 1920
aagacggagt gcgacgtgcg ggaagtggag ctgatgcggt gcctgcggct gcgtcagtcc 1980
tccctggagc ctgtggcctt ccggctgccc cgagtccgga aagagttctt ccaggatgac 2040
gtgttcccag acacggctgt gatctgggag cctgtgctca gtgccgaggc ctggctgcaa 2100
ggcgctaatg ggcagccctg gcttctcagc ctgcagcctc ctgacatgag cccagtgagc 2160
caagcccccc gagaggcccc tgctcgtcgg gccccatcct cagcgcagta cctggaagaa 2220
aagtctgacc agcaaaagaa ggaggagctg ctgaatgcca tggtggcaaa actggggaac 2280
cgggaggacc cactccccca ggactccttt gaaggcgtgg acgaggacga gtgggactag 2340
cctgcgcccc cgtcacctcc acctcacctg tgctgccact tcctagtgca cacctcacgg 2400
ctcatcctca agctggaaga tacctctctg gccccggcac atgtcacccc tgcactcctg 2460
ccttcccgtg ggcacttcca catcctctgg gcctctggca gttcccaggg actgttttca 2520
cctctgctgt ctctggggtc agctgctgct catcagctgc ccgctagcat gtggccaggg 2580
gtgcagggtg gcggggggtc agcagcatgt ccctgggcag gccctgggca ccctgtctcc 2640
cctggtctca ctgctgacct gggctggtcc cagcctggat tggcctcatc caggatcttt 2700
ggtcacccca cgctgcccca tcttgcctgc tgttccagtt ctggtcaagg gccttggggg 2760

ctggcccccc accaggcctt ctagagcagc accagtctca gggccctggg accagctgcc 2820
ctacttccca ggtttgtagc caggagaagg gggcatcaca gagctgatgg tccaataagg 2880
ggggtgtgag ccccgcaggg actggcccgc acctgccttg gatgttttca gcaattaaac 2940
tttttaagc tggcaaaaaa aaaaa 2965

<210> 87
<211> 2823
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2786701CB1

<400> 87
cggaggcagc ctagcctcgc gccccgcccg ttgcttctgc cctccggcct tcccgccgcc 60
gtcgccggga ccagccgctc ggggccgggc tgatacagcc gcttcaccgt gcccctgccc 120
gcgaccatgg cctcctccga ggtggcgcgg cacctgctct ttcagtctca catggcaacg 180
aaaacaactt gtatgtcttc acaaggatca gatgatgaac agataaaaag agaaaacatt 240
cgttcgttga ctatgtctgg ccatgttggt tttgagagtt tgcctgatca gctggtgaac 300
agatccattc agcaaggttt ctgctttaat attctctgtg tgggggaaac tggaattgga 360
aaatcaacac tgattgacac attgtttaat actaattttg aagactatga atcctcacat 420
ttttgcccaa atgttaaact taaagctcag acatatgaac tccaggaaag taatgttcaa 480
ttgaaattga ccattgtgaa tacagtggga tttggtgacc aaataaataa agaagagagc 540
taccaaccaa tagttgacta catagatgct cagtttgagg cctatctcca agaagaactg 600
aagattaagc gttctctctt tacctaccat gattctcgca tccatgtgtg tctctacttc 660
atttcaccga caggccactc tctgaagaca cttgatctct taaccatgaa gaaccttgac 720
agcaaggtaa acattatacc agtgattgcc aaagcagata cggtttctaa aactgaatta 780
cagaagttta agatcaagct catgagtgaa ttggtcagca atggcgtcca gatataccag 840
ttcccaacgg atgatgacac tattgctaag gtcaacgctg caatgaatgg acagttgccg 900
tttgctgttg tgggaagtat ggatgaggta aaagtcggaa acaagatggt caaagctcgc 960
cagtacccttt ggggtgttgt acaagtggaa aatgaaaacc actgtgactt tgtaaagctg 1020
cgggaaatgc tcatttgtac aaatatggag gacctgcgag agcagaccca taccaggcac 1080
tatgagcttt acaggcgctg caaactggag gaaatgggct ttacagatgt gggcccagaa 1140
aacaagccag tcagtgttca agagacctat gaagccaaaa gacatgagtt ccatggtgaa 1200
cgtcagagga aggaagaaga aatgaaacag atgtttgtgc agcgagtaaa ggagaaagaa 1260
gccatattga aagaagctga gagagagcta caggccaaat ttgagcacct taagagactt 1320
caccaagaag agagaatgaa gcttgaagaa aagagaagac ttttggaaga agaaataatt 1380
gctttctcta aaaagaaagc tacctccgag atatttcaca gccagtcctt tctggcaaca 1440
ggcagcaacc tgaggaagga caaggaccgt aagaactcca attttttgta aaacagaagt 1500
tccagagcac agaaggtcat catcacaagc aaactttatt aaaaaaaaac tagaagtgtg 1560
ctttgatttt gctgttattt gttttatcac ttctatattt ggtgaacagc cacagttact 1620
gatatttatg gaaaagtact ttcaagtaca aggtcaatac ataagccaga gtgaatgata 1680
ctacaagttg agcatctcta attcaaaaat ctgaaatcca gaagcttcaa aatctgaatc 1740
tttttgagca ctgacttgac cccacaagtg gaaaattccc cacccgacac ctttgctttc 1800
tgatggttca gtttaaacag attttgtttc ttgcacaaaa tttttgtata aattactttc 1860
aggctatatg tataaggtgg atgtgaaaca tgaattatgt aattagagtc gggtcccgtt 1920
gtgtatatgc agatattcca aacctgaaat ccaaaacact tctggtccct agcattttgg 1980
ataagggata ctcagcttgt acctatatat tcatatatat tcactgttgt tagaaatgtt 2040
taagttgctg ttctgtgatg aatctaaatc ttttctcttg ctaccaagct attgtcactg 2100
cagtgcatta taccaaagag cgaagtcagt gccactgaaa atacagaacc cattaatatc 2160
gtggctatct gattacattt atattccaag atgaaccttt tttatatatg ctaaaaattt 2220
tggggaatat gttttgggat gtattatgga gctaaaactc taacctctta atagttttat 2280
agaacttaaa aatttttttat acaattaccc aattggtgat atgatcttaa gcttttgtgt 2340
cagattattt aatatgatga cttcatgctt tattatgcct tattatggct gacgtattac 2400
tgtggtgaaa caaaatatct ttaaaagtta aaacatccag atatataagc tattttttcc 2460
taaggataaa gtacctttga gcatgagtgt atcacagctt tcattaggaa aacttttcat 2520
tacatacttg tttaaactct gtcttccagg gtaaaaataa taaggttgaa tcattttatt 2580
aaaaatactt tttaagaaaa taactatgaa catctgaata ttaaagatat aaaaatgcac 2640
ataattcata tttcaggtgg tatttgcatt cagtgcctta ctggtattct cagaacattt 2700
taatgatttc taacatttct taacagtcat agatatatac attttcattt tttgtacttg 2760
aatattctaa ataaaactga catttactct tgacaaataa aacatatatt tactaaaaaa 2820
aaa 2823

<210> 88
<211> 1549
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3068538CB1

<400> 88
gcagacccgg cacgcaggtg ggggccggcg gggtccgtgg ccagagctgc agagagacaa 60
ggcggcggcg gctgctgtgc tgggtgcagt gaggaagagg ccctcggtgg tgcccatggc 120
tggccaggat cctgcgctga gcacgagtca cccgttctac gacgtggcca gacatggcat 180
tctgcaggtg gcaggggatg accgctttgg aagacgtgtt gtcacgttca gctgctgccg 240
gatgccaccc tcccacgagc tggaccacca gcggctgctg gagtatttga agtacacact 300
ggaccaatac gttgagaacg attataccat cgtctatttc cactacgggc tgaacagccg 360
gaacaagcct tccctgggct ggctccagag cgcatacaag gagttcgata ggaagtacaa 420
gaagaacttg aaggccctct acgtggtgca ccccaccagc ttcatcaagg tcctgtggaa 480
catcttgaag cccctcatca gtcacaagtt tgggaagaaa gtcatctatt tcaactacct 540
gagtgagctc cacgaacacc ttaaatacga ccagctggtc atccctcccg aagttttgcg 600
gtacgatgag aagctccaga gcctgcacga gggccggacg ccgcctccca ccaagacacc 660
tccgccgcgg cccccgctgc ccacacagca gtttggcgtc agtctgcaat acctcaaaga 720
caaaaatcaa ggcgaactca tccccccctgt gctgaggttc acagtgacgt acctgagaga 780
gaaaggcctg cgcaccgagg gcctgttccg gagatccgcc agcgtgcaga ccgtccgcga 840
gatccagagg ctctacaacc aagggaagcc cgtgaacttt gacgactacg gggacattca 900
catccctgcc gtgatcctga agaccttcct gcgagagctg ccccagccgc ttctgacctt 960
ccaggcctac gagcagattc tcgggatcac ctgtgtggag agcagcctgc gtgtcactgg 1020
ctgccgccag atcttacgga gcctcccaga gcacaactac gtcgtcctcc gctacctcat 1080
gggcttcctg catgcggtgt cccgggagag catcttcaac aaaatgaaca gctctaacct 1140
ggcctgtgtc ttcgggctga atttgatctg gccatcccag ggggtctcct ccctgagtgc 1200
ccttgtgccc ctgaacatgt tcactgaact gctgatcgag tactatgaaa agatcttcag 1260
caccccggag gcacctgggg agcacggcct ggcaccatgg gaacagggga gcagggcagc 1320
ccctttgcag gaggctgtgc cacggacaca agccacgggc ctcaccaagc ctaccctacc 1380
tccgagtccc ctgatggcag ccagaagacg tctctagtgt tgcgaacact ctgtatattt 1440
cgagctacct cccacacctg tctgtgcact tgtatgtttt ataaacttgg catctgtaaa 1500
aataaccagc cattagatga attcagaacc ttctaatgaa aaaaaaaaa 1549

<210> 89
<211> 1722
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5159072CB1

<400> 89
gcaagaggga gccggcccga cgcggaccgc ttccctgcag tgccccgagt cccgggcccg 60
cgccgccgcc gcccggctcc gctcgcggcc cctctgtctg caggcgtgcc ccggcggcgg 120
cggagagccg tcctcggccg aggaggctgg gaaacgcgag cgcaggcggc agagaggcct 180
caacgccgtc cctttcgcca ccgccttttc cttgcctcgc gccgctgtgc atttctctcc 240
ttttcctttg ttttttggc ccctcgcggg tgtgggcatt gttggttagc aaaagtgcag 300
cctcaagatg gctgatggca acgaggatct gcgggctgac gacttgcctg ggccagcctt 360
cgagagctat gagtccatgg agcttgcctg ccccgctgag cgcagcggcc acgtagccgt 420
cagcgacggg cgccacatgt tcgtctgggg cggctacaag agtaatcaag tcagaggatt 480
atatgacttt tatctgccta gagaagaact atggatctac aacatggaga ctggaagatg 540
gaaaaaaatc aacactgaag gtgatgttcc tccttctatg tcaggaagct gtgctgtgtg 600
tgtagacagg gtgctgtact tgtttggagg acaccattca agaggcaata ccaataagtt 660
ctacatgctg gattcaaggt ctacagacag agtgttacag tgggaaagaa ttgattgcca 720
aggaattcct ccatcatcaa aggacaaact tggtgtctgg gtatataaaa acaagttaat 780
attttttgga gggtatggat atttgcctga agataaagta ttgggaactt ttgaattcga 840
tgaaacatct ttttggaatt caagtcatcc aagaggatgg aatgatcatg tacatatttt 900
agatactgaa acatttacct ggagccagcc tataactact ggtaaagcac cttcacctcg 960

tgctgcccat gcctgtgcaa ctgtcggaaa tagaggcttc gtgtttggag gcagatatcg 1020
agatgctaga atgaatgatc ttcactatct taatctggat acatgggagt ggaatgaatt 1080
aattccacaa ggcatatgcc cagttggtcg atcttggcac tcactaacac cagtttcttc 1140
agatcatctt tttctctttg gaggatttac cactgataaa cagccactaa gtgatgcctg 1200
gacttactgc atcagtaaaa atgaatggat acaatttaat catccatata ccgaaaaacc 1260
aaggttatgg cacacagctt gtgccagcga tgaaggagaa gtaattgttt ttggtggatg 1320
tgccaacaac ttgcttgtcc atcacagagc tgcacacagt aatgaaatac taatattttc 1380
agttcaacca aaatctcttg tacggctaag cttagaagca gtcatttgct ttaaagaaat 1440
gttagccaac tcatggaact gccttccaaa acacttactt cacagtgtta atcagaggtt 1500
tggtagtaac aacacttctg gatcttaagg cttcataaat aatgcctatg atcaccttgc 1560
atggacagca atcctgtaaa catcacagag tggcatcatt tgtataatta tatgcattgt 1620
tgtagtttgc acctgttggt tttaatgtgc atgtgaatgg cctagagaac ctatttttgt 1680
gtctaaagtt tacaataaat gtatttaaca ccaaaaaaaa aa 1722

<210> 90
<211> 1264
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5519057CB1

<400> 90
agcgcgcgct cttgcggtgg cgtaatctct cagcctttct gtgtctcctt tcctccgcct 60
cagtttggtg cgggtcgggg gaatggctga ggagatggag tcgtcgctcg aggcaagctt 120
ttcgtccagc ggggcagtgt caggggcctc agggtttttg cctcctgccc gctcccgcat 180
cttcaagata atcgtgatcg gcgactccaa tgtgggcaag acatgcctga cctaccgctt 240
ctgcgctggc cgcttccccg accgcaccga ggccacgata ggggtggatt tccgagaacg 300
agcggtggag attgatgggg agcgcatcaa gatccagcta tgggacacag caggacaaga 360
acgattcaga aagagcatgg ttcagcacta ctacagaaat gtacatgctg ttgtcttcgt 420
gtatgatatg accaacatgg ctagttttca tagcctacca tcttggatag aagaatgcaa 480
acaacatttg ctagccaatg atataccacg gattcttgtt ggaaataaat gtgacttgag 540
aagtgccata caggtaccca cagacttggc acaaaaattt gctgacacac acagtatgcc 600
tttgtttgaa acgtctgcta aaaaccccaa tgataatgac catgtggaag ctatatttat 660
gaccttggct cataagctta agtgccacaa accattaatg cttagtcagc cccctgataa 720
tggaattatc ctgaagcctg aaccaaagcc tgcaatgacg tgctggtgct aaataacagt 780
ctttattata ttatctaatt ttgactaaag aaatactttt gaagtatgac agtattaagt 840
cataagattt aatctcaact ataatgggtc atcttgacac tttgctgttt gtcattgtca 900
cgcttttgta ttttgtatct acttaagttt gtcactgtga caacacagga aaagttggtt 960
ttcaggtgag attgaaaatg aagcaaagat aggatgaatc tgaacatctc tccatctaga 1020
gcccaatgaa ggaagcttca aatgagaaca tgatggaatc agtaaccatt caatcttttg 1080
tcctaggatt ggaaaaaaat gttaaaggtt taggacacac ctaatagtat gtcctttgaa 1140
tgggaagttt tcttaatagg ataaaaactg gtatttcctt ctccccagag tacttttttg 1200
ttttttccat agagacgggg tcttgctatg ttgtccaggc tggccttgag ctcctgggct 1260
caag 1264

<210> 91
<211> 2640
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID.No: 035379CB1

<400> 91
cggccgaggg ggcatcatga agcgggctgg cggcgctgcg tcccgggcgc gcgcgggcgg 60
gaggtgcttc ccaaggaccg tagatgcctc tctagagcat gagctcaggc aagagtgccc 120
gctacaaccg cttctccggg gggcccagca atcttcccac cccagacgtc accacaggga 180
ccagaatgga aacgaccttc ggacccgcct tttcagccgt caccaccatc acaaaagctg 240
acgggaccag cacctacaag cagcactgca ggacaccctc ctcctccagc acccttgcct 300
actccccgcg ggacgaggag gacagcatgc cccccatcag cactccccgc cgctccgact 360

ccgccatctc tgtccgctcc ctgcactcag agtccagcat gtctctgcgc tccacattct 420
cactgcccga ggaggaggag gagccggagc cactggtgtt tgcggagcag ccctcggtga 480
agctgtgctg tcagctctgc tgcagcgtct tcaaagaccc cgtgatcacc acgtgtgggc 540
acacgttctg taggagatgc gccttgaagt cagagaagtg tcccgtggac aacgtcaaac 600
tgaccgtggt ggtgaacaac atcgcggtgg ccgagcagat cggggagctc ttcatccact 660
gccggcacgg ctgccgggta gcgggcagcg ggaagccccc catctttgag gtggaccccc 720
gagggtgccc cttcaccatc aagctcagcg cccggaagga ccacgagggc agctgtgact 780
acaggcctgt gcggtgtccc aacaacccca gctgcccccc gctgctcagg atgaacctgg 840
aggcccacct caaggagtgc gagcacatca aatgccccca ctccaagtac gggtgcacgt 900
tcatcgggaa ccaggacact tacgagaccc acctggagac ttgccgcttc gagggcctga 960
aggagtttct gcagcagacg gatgaccgct tccacgagat gcacgtggct ctggcccaga 1020
aggaccagga gatcgccttc ctgcgctcca tgctgggaaa gctctcggag aagatcgacc 1080
agctagagaa gagcctggag ctcaagtttg acgtcctgga cgaaaaccag agcaagctca 1140
gcgaggacct catggagttc cggcgggacg catccatgtt aaatgacgag ctgtcccaca 1200
tcaacgcgcg gctgaacatg ggcatcctag gctcctacga ccctcagcag atcttcaagt 1260
gcaaagggac ctttgtgggc caccagggcc ctgtgtggtg tctctgcgtc tactccatgg 1320
gtgacctgct cttcagtggc tcctctgaca agaccatcaa ggtgtgggac acatgtacca 1380
cctacaagtg tcagaagaca ctggagggcc atgatggcat cgtgctggct ctctgcatcc 1440
aggggtgcaa actctacagc ggctctgcag actgcaccat cattgtgtgg gacatccaga 1500
acctgcagaa ggtgaacacc atccgggccc atgacaaccc ggtgtgcacg ctggtctcct 1560
cacacaacgt gctcttcagc ggctccctga aggccatcaa ggtctgggac atcgtgggca 1620
ctgagctgaa gttgaagaag gagctcacag gcctcaacca ctgggtgcgg gccctggtgg 1680
ctgcccagag ctacctgtac agcggctcct accagacaat caagatctgg gacatccgaa 1740
cccttgactg catccacgtc ctgcagacgt ctggtggcag cgtctactcc attgctgtga 1800
caaatcacca cattgtctgt ggcacctacg agaacctcat ccacgtgtgg gacattgagt 1860
ccaaggagca ggtgcggacc ctcacgggcc acgtgggcac cgtgtatgcc ctggcggtca 1920
tctcgacgcc agaccagacc aaagtcttca gtgcatccta cgaccggtcc ctcagggtct 1980
ggagtatgga caacatgatc tgcacgcaga ccctgctgcg tcaccagagc agtgtcaccg 2040
cgctggctgt gtcccggggc cgactcttct caggggctgt ggatagcact gtgaaggttt 2100
ggacttgcta acaggatcca ggccaggctg tggtttcccc tgaaccagcc ctggaccttt 2160
ctgagccagg ctggccacat ggggtggtct cggggtttct gcctgccccg tgggcatagg 2220
tggacaggct ctggcagccg ggcagtgccc tccccgtccc atgctcggcg agcctccctc 2280
tactcggcac tgtccttgct gcccagcccc tctctgggtg ccaggtacga cgcttgcccc 2340
ggcccacccct ccatccccac cctccatccc caccctagat ggagcgaggg ccttttact 2400
caccttttct accgttttta gactgtatgt agattggtta cctcctggtt gaaataaatg 2460
ctccacagac tgtggctgtg agtggggaca gctcctcggg acaagggggc tgtgtgtggc 2520
cttgaggttg gtgtgcacag gcactggctg ctgtgagtgg gggggcatgg ggcagtttcc 2580
tttggtggac cccaggactt cggcccactc cggggcctcc cctccctgct aggaggtaac 2640

<210> 92
<211> 2071
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 275354CB1

<400> 92
gtgggcggaa ctcctagcgg acacctcgtg gagtccggcc ggaagagcaa ccgagatgaa 60
ggtgaagatg ctgagccgga atccggacaa ttatgtccgc gaaaccaagt tggacttaca 120
gagagttcca agaaactatg atcctgcttt acatcctttt gaggtcccac gagaatatgt 180
aagagcttta aatgctacca aactggaacg agtatttgca aaaccattcc ttgcttcgct 240
ggatggtcac cgtgatggag tcaattgctt ggcaaagcat ccagagaagc tggctactgt 300
cctttctggg gcgtgtgatg gagaggttag aatttggaat ctaactcagc ggaattgtat 360
ccgtacaata caagcacatg aaggctttgt acgaggaata tgtactcgct tttgtgggac 420
ttcttttttc actgttggtg atgacaaaac tgtgaagcag tggaaaatgg atgggccagg 480
ctatggagac gaggaagagc cattacatac aatattagga aagacagtgt atactgggat 540
tgatcatcac tggaaagaag ctgtttttgc cacatgtgga cagcaagtag acatttggga 600
tgaacaaaga actaatccta tatgttcaat gacctgggga tttgacagta taagtagtgt 660
taaatttaac ccaattgaga catttctctt gggaagttgt gcatctgaca ggaatatagt 720
actgtacgat atgaggcaag ctactccttt gaaaaaggtt atcttagata tgagaacaaa 780
tacaatctgt tggaacccta tggaagcttt catttttaca gcagcaaatg aagattataa 840

cttatatact tttgatatgc gtgcactgga cactcctgta atggtccata tggatcatgt 900
atctgcagtg cttgatgtgg attactctcc cactgggaag gagtttgtgt ctgctagttt 960
cgataaatct attcgaatct ttcctgtaga caaaagtcga agcagggagg tatatcatac 1020
aaagagaatg caacatgtta tctgtgtaaa atggacttct gacagcaagt atattatgtg 1080
tggatctgat gaaatgaaca ttcgcctgtg gaaagctaat gcttctgaaa aattgggtgt 1140
gcttacatca cgagaaaaag cagccaagga ttataaccag aaattgaagg agaaatttca 1200
gcattatcct catataaaac gtatagctcg tcatcgacat ctaccaaaat ctatctatag 1260
ccagattcag gaacagcgca tcatgaaaga agctcgtcga cgaaaggaag tgaatcgtat 1320
taaacacagc aagcctggat ctgtgccact tgtgtcagag aagaagaaac acgtagtggc 1380
agttgtaaaa taattggtat tcctaacaat cctgatgtat aattatttgt tacttttgat 1440
ttgagaactc tacaaataaa agtgctggga ctagattaat tgcaaacatt ttagttatat 1500
gtgtagagct ttattgttac tccttttagc taccctgaaa aatgatcctt aaaggtggcc 1560
tagttggtaa gactgtttta tccttaatct gcattcttct ttcattgtag aatacagtat 1620
ttgcaactca tttttcttg tttttattac agatatactt actttctctt tgatctatta 1680
ttgtagacac tatacattca aattgacatt taagaccaaa catctcttat gttatcttta 1740
atattacttt gaataatgat tgcaatgatg tttcttcctg tgattccaca taacatttag 1800
aataatgatg tcaattttttt acaactgaat ttatttctag tgctttactt atatttggct 1860
ttttgactct tttaaaacaa tcagcctgca tttatataac ttttataaat aataatataa 1920
tttgggtcaa gttaagatat taaaagttcc tttcagcatt gaaactttgg cctatttttg 1980
gtaaataatt ttcaatctca ctaaatccta aatagctctg tgtaacatag gtttttcttt 2040
ttttaatcat aaacttaata aactttgtgg a 2071

<210> 93
<211> 2149
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 311658CB1

<400> 93
cattattttt aaaaatatta cccactcttg atagtgtatc tgcactgaga cacgtactgg 60
aagctatata ttgtttgaca tcccaatcta aaacaactca tctttcttac ttatgactag 120
agttcctcct cttcatttat attcttttct tggtgaacat cagtgtctac caatttctaa 180
atgcaaagga gaaagataca atttaagcg aaatggtggt gatatgcaca acttgcagaa 240
ggttacataa aacttgggtt ttcagagatg attttttctc ttctttttag gatatgttca 300
aggaatgagt gatttacttt cccctctttt atatgtgatg gaaaatgaag tggatgcctt 360
ttggtgcttt gcctcttaca tggaccaaat gcatcagaat tttgaagaac aaatgcaagg 420
catgaagacc cagctaattc agctgagtac cttacttcga ttgttagaca gtggattttg 480
cagttactta gaatctcagg actctggata cctttatttt tgcttcaggt ggcttttaat 540
cagattcaaa agggaattta gttttctaga tattcttcga ttatgggagg taatgtggac 600
cgaactacca tgtacaaatt tccatcttct tctctgttgt gctattctgg aatcagaaaa 660
gcagcaaata atggaaaagc attatggctt caatgaaata cttaagcata tcaatgaatt 720
gtccatgaaa attgatgtgg aagatatact ctgcaaggca gaagcaattt ctctacagat 780
ggtaaaatgc aaggaattgc cacaagcagt ctgtgagatc cttgggcttc aaggcagtga 840
agttacaaca ccagattcag acgttggtga agacgaaaat gttgtcatga ctccttgtcc 900
tacatctgca tttcaaagta atgccttgcc tacactctct gccagtggag ccagaaatga 960
cagcccaaca cagataccag tgtcctcaga tgtctgcaga ttaacacctg catgatcact 1020
gttcttgctt ttttgggaag agacactttg ttgcaaccct ttttcaagta cttgaaagtt 1080
gaaaatttga aatcttggta ttgatcatgc tttaaggttt atgtaaagaa agtgtactga 1140
tgttcttaca ttaaagcttt acaaagattt aaactaatta tttttgtagt tacttctacc 1200
aaatagcctt tccttttcga taacattcct cagtattttt atagccaagt acattttatt 1260
ttcttgctga tgaactggaa ttggataaat attgcaagtg gatgagttgg aaattatgca 1320
ctttgaaaaa cattcacttt gtttaagctt attgggtttc agatttgatt aaattaaatg 1380
tggaggcttt ctatagcatt ctaagctgag aagtagattg ttacccagta atgaaataaa 1440
aaataaaaat aaaaggattt ttttctctat tgtttacgac agtactcagc ttaaatattt 1500
atgctggtca aatgtgattt aaattggaca ttttcatcaa tgcagtctaa tgtgtagata 1560
aatatttcaa ccataataag tggattggca gtatattttt tacattgaac ttttcttcac 1620
ttgtatataa agattatata taagtactta tttatgagta taagaaaggt taggcatatt 1680
ttcattaact gaataaacga cttgatttat ataacctggt ttatcaaaat ttaacatggc 1740
ttcagtatga gatcttttttc aaaactattt tcttaaacat ttatttcatg agattatgtt 1800
caaccctgta cctggtgtaa ttttaaaatt aattgcttgt aacctcactt tactaataat 1860

gtttattatc tttcctaata atgcattaac tgattaatca ggtgtttaaa tttttataaa 1920
atactcttgc aaaaagttta tttgaaaaat ttctagatgg tctcatgagt ttcaaaataa 1980
taatttttgt gtatgaacaa agctgttgtt tttaccatgc agtattgcat gattttaagt 2040
tatgtggaat taacataact gattttgttt taattgtaag ttgttaactc ctgtatatat 2100
cattaaaata aatctgaagt tgaagtagtg tttttagtta aattatact 2149

<210> 94
<211> 2332
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1251632CB1

<400> 94
gccaccccag aactgggcag cagcctcaag aagaagaagc ggctctccca gtcagatgag 60
gatgtcatta ggctaatagg acagcacttg aatggcttag ggctcaacca gactgttgat 120
ctcctcatgc aagagtcagg atgtcgttta gaacatcctt ctgctaccaa attccgaaat 180
catgtcatgg aaggagactg ggataaggca gaaaatgacc tgaatgaact aaagccttta 240
gtgcattctc ctcatgctat tgtgaggatg aagtttttgc tgctgcagca gaagtaccta 300
gaatacctgg aggatggcaa ggtcctggag gcacttcaag ttctacgctg tgaattgacg 360
ccgctgaaat acaatacaga gcgcattcat gttcttagtg ggtatctgat gtgtagccat 420
gcagaagacc tacgtgcaaa agcagaatgg gaaggcaaag ggacagcttc ccgatctaaa 480
ctattggata aacttcagac ctatttacca ccatcagtga tgcttccccc acggcgttta 540
cagactctcc tgcggcaggc ggtggaacta caaagggatc ggtgcctata tcacaatacc 600
aaacttgata ataatctaga ttctgtgtct ctgcttatag accatgtttg tagtaggagg 660
cagttcccat gttatacgca gcagatactt acggagcatt gtaatgaagt gtggttctgt 720
aaattctcta atgatggcac taaactagca acaggatcaa aagatacaac agttatcata 780
tggcaagttg atccggatac acacctgcta aaactgctta aaacattaga aggacatgct 840
tatggcgttt cttatattgc atggagtcca gatgacaact atcttgttgc ttgtggccca 900
gatgactgct ctgagctttg gctttggaat gtacaaacag gagaactaag gacaaaaatg 960
agccagtctc atgaagacag tttgacaagt gtggcttgga atccagatgg gaagcgcttt 1020
gtgactggag gtcagcgtgg gcagttctat cagtgtgact tagatggtaa tctccttgac 1080
tcctgggaag ggtaagagt gcaatgcctt tggtgcttga gtgatggaaa gactgttctg 1140
gcatcagata cacaccagcg aattcggggc tataacttcg aggaccttac agataggaac 1200
atagtacaag aagatcatcc tattatgtct tttactattt caaaaaatgg ccgattagct 1260
ttgttaaatg tagcaactca gggagttcat ttatgggact tgcaagacag agttttagta 1320
agaaagtatc aaggtgttac acaagggttt tatacaattc attcatgttt tggaggccat 1380
aatgaagact tcatcgctag tggcagtgaa gatcacaagg tttacatctg gcacaaacgt 1440
agtgaactgc caattgcgga gctgacaggg cacacacgta cagtaaactg tgtgagctgg 1500
aacccacaga ttccatccat gatggccagc gcctcagatg atggcactgt tagaatatgg 1560
ggaccagcac cttttataga ccaccagaat attgaagagg aatgcagtag catggatagt 1620
tgatggtgaa tttggagcag acgacttctg tttaacttaa aattagtcgt attttaatgg 1680
cttgggattt ggtgcaaaca aacatgattg atagctggac agacatgctc gtcatgaaaa 1740
aagaaccatt tctgaagccc gattggggcc aaacatttac accttgcttc atagtaacca 1800
gttgagatga agcacgtcgt tagaacgttg ttggacacca tgttgaatta ttcccccatc 1860
ggttgtgaag aactgtgcta cattcaggct tacccattga actcagtata tatatttttt 1920
ttccttcctg tcttttgtct ggcaggatac cattcttgtt gctcttctgt gtaatgaagt 1980
ttaaatgctt gtttggaaaa ctttatttaa cagtttagaa ggcttgatag aaagagtgca 2040
ttagtctgaa gagtatacat tggataggaa agaatttcct tcttttgttt ctccaaatct 2100
ttccgcctta tttagcttga gatctttgca gcttggttca tggattctag ccttgcccgt 2160
tgcgcagtat atactgatcc agatgataaa ccagtgaact atgtcaaaag cactctcaat 2220
attacatttg acaaaaagtt ttgtactttt cacatagctt gttgccccgt aaaagggtta 2280
acagcacaat tttttaaaaa taaattaaga agtatttata ggaaaaaaaa aa 2332

<210> 95
<211> 1751
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 1331955CB1

<400> 95
gcccgaatcg actccggaga caacgccggg cgacgccacc tgcgcaggtc ccggaggccg 60
ctggtgtctg tgtacagggc gtgctgtctg tggaaacgcg agggcacact agcactttcc 120
tggaaggacc ccagacccac caagccactc agtccctgga cgagtttcca ctcaccccga 180
ctgcctctgt caccgggtcc ctccaccctt gtctcctgtg cggccagcgt cagagccatg 240
gcgacggagg agaagaagcc cgagaccgag gccgccagag cacagccaac cccttcgtca 300
tccgccactc agagcaagcc tacacctgtg aagccaaact atgctctaaa gttcaccctt 360
gctggccaca ccaaagcagt gtcctccgtg aaattcagcc cgaatggaga gtggctggca 420
agttcatctg ctgataaact tattaaaatt tggggcgcgt atgatgggaa atttgagaaa 480
accatatctg gtcacaagct gggaatatcc gatgtagcct ggtcgtcaga ttctaacctt 540
cttgtttctg cctcagatga caaaaccttg aagatatggg acgtgagctc gggcaagtgt 600
ctgaaaaccc tgaagggaca cagtaattat gtcttttgct gcaacttcaa tccccagtcc 660
aaccttattg tctcaggatc ctttgacgaa agcgtgagga tatgggatgt gaaaacaggg 720
aagtgcctca agactttgcc agctcactcg gatccagtct cggccgttca ttttaatcgt 780
gatggatcct tgatagtttc aagtagctat gatggtctct gtcgcatctg ggacaccgcc 840
tcaggccagt gcctgaagac gctcatcgat gacgacaacc cccccgtgtc ttttgtgaag 900
ttctccccga acggcaaata catcctggcc gccacgctgg acaacactct gaagctctgg 960
gactacagca aggggaagtg cctgaagacg tacactggcc acaagaatga gaaatactgc 1020
atatttgcca atttctctgt tactggtggg aagtggattg tgtctggctc agaggataac 1080
cttgtttaca tctggaacct tcagacgaaa gagattgtac agaaactaca aggccacaca 1140
gatgtcgtga tctcaacagc ttgtcaccca acagaaaaca tcatcgcctc tgctgcgcta 1200
gaaaatgaca aaacaattaa actgtggaag agtgactgct aagtcccttt gctcctgccc 1260
gcgagagact gtcgggaagt tgacccggat tggcaagaaa cagggtgtct tggaggtggt 1320
cccccagatc tgcgcctggg ggtcaggaca gggcctgatt tgagcctcct ctctgaagat 1380
gatttggccg agcggaaggt gtggaccacc ggaaagttct taaaagttgc tggtgacatt 1440
tcttgccaat tctaacactg tctagggaag agttcctagt ctattgtgtt caaacagagt 1500
caacaaaagt ttttaatttt ttattacaga agggtgaagt tcaatttaac atgcgttgtg 1560
tttttcagt aaacgttctg tatctttttg atattccatg acccagtgca cgctgtggcc 1620
tgtcaccgcc accgtggccc cgccagctgg cctccccttt ggcccacgcc ggccgccccc 1680
attctctgct gcgtagatgc cctggcccag ggcctgactc tccattcccg ccagtagggg 1740
taccgagctc g 1751

<210> 96
<211> 1285
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1412614CB1

<400> 96
cggagaaaaa gctgacctaa tgaaactgtg gcaacgtcag cgcttgaggc ttgaagaggg 60
agacaagcta aaagaggata tccagttgtt tcatgatggc atttgcacct ccaaaaaaca 120
cagatggtcc caaaatgcag acaaagatga gcacctggac acccctaaac catcagctat 180
tgaatgaccg ggtatttgaa gaaagaagag ccctgcttgg caaatggttt gacaaatgga 240
cagactctca aagaagaaga atcctcacag gcctgttgga gcgctgctcg ctgtcccagc 300
aaaagttctg ctgtcgaaag cttcaagaga aaattccagc agaagccctg gactttacaa 360
ccaagcttcc aagggtgtta tctttataca tcttttcttt cctggaccct cggagccttt 420
gtcgttgtgc acaggtgtgc tggcattgga agaaccttgc tgagctggac cagctctgga 480
tgctgaaatg tttacggttt aactggtaca tcaatttctc tccaactccc tttgagcagg 540
ggatctggaa gaagcactat attcaaatgg tgaaagaact tcatattacc aagcctaaga 600
cacccccaaa ggatggattt gtaatcgctg acgttcaact agttacaagc aattctccag 660
aggaaaaaca gtcccccttta tcagcttttc ggtcctcttc ctctttaaga aagaagaata 720
actcagggga gaaagcactt ccaccctggc gatcttctga taagcaccca acagatatca 780
ttcgttttaa ttacctagac aaccgtgacc ccatggagac tgtccagcaa ggaagaagaa 840
aaagaaacca aataacccca gacttcagcc gacagtcaca tgataagaaa aataaattgc 900
aggacagaac taggctaaga aaagcacaat caatgatgtc gaggagaaat cccttcccac 960
tatgtcccta agtgccagct ctcccctaaa agttccagct catctcgcct ggcctccccc 1020
tgagtcagtg ggactcccag ccactgccac cacagctgaa attctcatgc agcatcctca 1080
caggcaccct gggccccaag catgactcat ccaggttcca gagccaaagt ggactgaaca 1140

tgggaagact tttattatag aaatgacaag atgctttgca cagtggagag ctgaatttac 1200
ttggctccca ttagaaactc tttcagctta agtacttatt gtggtagtga gtcctacggt 1260
atttcagtaa aaaggaattc atggc 1285

<210> 97
<211> 3260
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1750781CB1

<400> 97
ccggaagacc gtcccggatg gcctcgggga ctgccagtgt gtggaggtga gctccgggat 60
tgccggcatt cccgcttctg ctggttgctt catgctgcag gctgcggccg tcagccctcg 120
ctcgcattgg tggcgctgag gtgccggggc agcaagtgac atgtcgtcgg gcctccgcgc 180
cgctgacttc ccccgctgga agcgccacat ctcggagcaa ctgaggcgcc gggaccggct 240
gcagagacag gcgttcgagg agatcatcct gcagtataac aaattgctgg aaaagtcaga 300
tcttcattca gtgttggccc agaaactaca ggctgaaaag catgacgtac caaacaggca 360
cgagataagt cccggacatg atggcacatg gaatgacaat cagctacaag aaatggccca 420
actgaggatt aagcaccaag aggaactgac tgaattacac aagaaacgtg gggagttagc 480
tcaactggtg attgacctga ataaccaaat gcagcggaag gacagggaga tgcagatgaa 540
tgaagcaaaa attgcagaat gtttgcagac tatctctgac ctggagacgg agtgcctaga 600
cctgcgcact aagctttgtg accttgaaag agccaaccag accctgaagg atgaatatga 660
tgccctgcag atcactttta ctgccttgga gggaaaactg aggaaaacta cggaagagaa 720
ccaggagctg gtcaccagat ggatggctga gaaagcccag gaagccaatc ggcttaatgc 780
agagaatgaa aaagactcca ggaggcggca agcccggctg cagaaagagc ttgcagaagc 840
agcaaaggaa cctctaccag tcgaacagga tgatgacatt gaggtcattg tggatgaaac 900
ttctgatcac acagaagaga cctctcctgt gcgagccatc agcagagcag ccacgagacg 960
ctctgtctct tccttcccag tcccccagga caatgtggat actcatcctg gttctggtaa 1020
agaagtgagg gtaccagcta ctgccttgtg tgtcttcgat gcacatgatg ggaagtcaa 1080
cgctgtgcag ttcagtccag gttcccggtt actggccact ggaggcatgg accgcagggt 1140
taagctttgg gaagtatttg gagaaaaatg tgagttcaag ggttccctat ctggcagtaa 1200
tgcaggaatt acaagcattg aatttgatag tgctggatct tacctcttag cagcttcaaa 1260
tgattttgca agccgaatct ggactgtgga tgattatcga ttacggcaca cactcacggg 1320
acacagtggg aaagtgctgt ctgctaagtt cctgctggac aatgcgcgga ttgtctcagg 1380
aagtcacgac cggactctca aactctggga tctacgcagc aaagtctgca taaagacagt 1440
gtttgcagga tccagttgca atgatattgt ctgcacagag caatgtgtaa tgagtggaca 1500
ttttgacaag aaaattcgtt tctgggacat tcgatcagag agcatagttc gagagatgga 1560
gctgttggga aagattactg ccctggactt aaacccagaa aggactgagc tcctgagctg 1620
ctcccgtgat gacttgctaa aagttattga tctccgaaca aatgctatca agcagacatt 1680
cagtgcacct gggttcaagt gcggctctga ctggaccaga gttgtcttca gccctgatgg 1740
cagttacgtg gcggcaggct ctgctgaggg ctctctgtat atctggagtg tgctcacagg 1800
gaaagtggaa aaggttcttt caaagcagca cagctcatcc atcaatgcgg tggcgtggtc 1860
gccctctggc tcgcacgttg tcagtgtgga caaaggatgc aaagctgtgc tgtgggcaca 1920
gtactgacgg ggctctcagg gctgggagga ccccagtgcc ctcctcagaa gaagcacatg 1980
ggctcctgca gccctgtcct ggcaggtgat gtgctgggta tagcatggac ctcccagaga 2040
agctcaagct atgtggcact gtagctttgc cgtgaatggg atttctgaag atttgactga 2100
ggtctctctt ggcctggaag aataacactg aaaaaacctg acgctgcggt cacttagcag 2160
aggctcaggt tcttgccttg ggaaacacta ctagctctga ccttccatac ctcacttggg 2220
ggagcacagg gccccgctgg gcctcctcac caacggcagt gccaaaatca gcccccacat 2280
caaggtggtg ttctctgtgc tttctctcgt ccttccaaag tcggttctgg cctaacgcat 2340
gtcccaacac cttgggttca tttgcccggt gaactcactt taagcattgg attaacggaa 2400
actcccgaac tacagacccc tccctggtgg gttgcatgaa tgtgtctcat tactgctgaa 2460
atgtcctcac atctcttcca ctgttcttca gagctttctg gctctctttc cccacaaaat 2520
tcgacacatt taaaaatctc cgtgtggctt taaaaaatgg ttttttgttt ttttgttttt 2580
ttgaggtggg agaggatgtg tgaaaatctt ttccagggaa atgggttcgc tgcagaggta 2640
aggatgtgtt cctgtatcga tctgcagaca cccagaaggt gggtgcacac tgcatgcttg 2700
ggggtgccaa gggattcgag acctccaaca tacttgtctg aaggtggtga ttctggccat 2760
ggcccctctg ccaagcctgt gtgcgatgcc cttggtgctt tagtgcaaga agcctaggct 2820
cagaagcaca gcagcgccat ctttccgttt caggggttgt gatgaaggcc aaggaaaaac 2880
atttatcttt actattttac ctacgtataa agttttagtt cattgggtgt gcgaaacacc 2940

cttttatca cttttaaatt tgcactttat tttttttctt ccatgcttgt tctctggaca 3000
tttggggatg tgagtgttag agctggtgag agaggagtca ggcggccttc ccaccgatgg 3060
tcctggcctc cacctgccct ctcttccctg cctgatcacc gctttccaat ttgcccttca 3120
gagaacttaa gtcaaggaga gttgaaattc acaggccagg gcacatcttt tatttatttc 3180
attatgttgg ccaacagaac ttgattgtaa ataataataa agaaatctgt tatatacttt 3240
tcaaaaaaaa aaaaaaaaaa 3260

<210> 98
<211> 1276
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1821658CB1

<400> 98
gcggcacccc caaggaagac cagcctgcct ctggtcggtt cctggcgctc tgcgtttcgt 60
gaccttgtcc agtagaaggc tatttaattt tcacaactgc ttgaattttg acatacaaga 120
tgaagcaaga tgcctcaaga aatgctgcct acactgtgga ttgtgaagat tatgtgcatg 180
tggtagaatt taatcccttt gagaatgggg attcaggaaa cctaattgca tatggtggca 240
ataattatgt ggtcattggc acgtgtacgt ttcaggaaga agaagcagac gttgaaggca 300
ttcagtataa aacacttcga acatttcacc atggagtcag ggttgatggc atagcttgga 360
gcccagagac tagacttgat tcattgcctc cagtaatcaa attttgtact tcagctgctg 420
atatgaaaat tagattattt acttcagatc ttcaggataa aaatgaatat aaggttttag 480
agggccatac cgatttcatt aatggtttgg tgtttgatcc caaagaaggc caagaaattg 540
caagtgtgag tgacgatcac acctgcagga tttggaactt ggaaggagtg caaacagctc 600
attttgttct tcattctcct ggcatgagtg tgtgctggca tcctgaggag acttttaagc 660
taatggttgc agagaagaat ggaacaatcc ggttttatga tcttttggcc caacaggcta 720
ttttatctct tgaatcagaa caagtgccat taatgtcagc acactggtgc ttaaaaaaca 780
ccttcaaagt tggagccgtt gcaggaaatg attggttaat ttgggatatt actcggtcca 840
gttatcctca aaataagaga cctgttcaca tggatcgagc ctgcttattc aggtggtcca 900
caattagtga aaatctgttt gcaaccactg gttatcctgg caaaatggca agccagtttc 960
aaattcatca tttaggacac cctcagccca tcctcatggg ttctgtagcc gttggatctg 1020
gactgtcctg gcatcgaact ctccctctgt gtgtaattgg aggagaccac aagctgttgt 1080
tttgggtgac tgaagtataa agtgttttct gtaccttaga ttcacaaact ttgtattttt 1140
agtacatatt ttgaagaatt tctatagtac atattttgaa gaatttttat atcaaatata 1200
ccgtatactt tagaaaatgt ctcagttgct tttattaaat aaaatgttga tggtttgaaa 1260
aattaaaaaa aaaaaa 1276

<210> 99
<211> 3608
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1872574CB1

<400> 99
gttttggttc tagccgctcg ccgtccttgc aggctctgcc gtcggaaagc cgctcattct 60
cgcttcccct tccctttccc ggctcaagtc cttcctctct ctttcctttc tttccgccta 120
tctttttttct gctgccgctc cgggtccggg ccattttccg ggccgggcgc actaaggtgc 180
gcggccccgg ggcccagtat atgacccgcc gtcctgctat ccttcgcttc ccccgcccca 240
tgtggctgcg gggccgcggc ggcgctgccc actatggccc ggaaagtagt tagcaggaag 300
cggaaagcgc ccgcctcgcc gggagctggg agcgacgctc agggcccgca gtttggctgg 360
gatcactcgc ttcacaaaag gaaaagactt cctcctgtga agagatcctt agtatactac 420
ttgaagaacc gggaagtcag gctacagaat gaaaccagct actctcgagt gttgcatggt 480
tatgcagcac agcaacttcc cagtctcctg aaggagagag agtttcacct tgggaccctt 540
aataaagtgt ttgcatctca gtggttgaat cataggcaag tggtgtgtgg cacaaaatgc 600
aacacgctat ttgtcgtaga tgtccagaca agccagatca ccaagatccc cattctgaaa 660
gaccgggagc ctggaggtgt gacccagcag ggctgtggta tccatgccat cgagctgaat 720
ccttctagaa cactgctagc cactggagga gacaacccca acagtcttgc catctatcga 780

ctacctacgc tggatcctgt gtgtgtagga gatgatggac acaaggactg gatcttttcc 840
atcgcatgga tcagcgacac tatggcagtg tctggctcac gtgatggttc tatgggactc 900
tgggaggtga cagatgatgt tttgaccaaa agtgatgcga gacacaatgt gtcacgggtc 960
cctgtgtatg cacacatcac tcacaaggcc ttaaaggaca tccccaaaga agacacaaac 1020
cctgacaact gcaaggttcg ggctctggcc ttcaacaaca agaacaagga actgggagca 1080
gtgtctctgg atggctactt tcatctctgg aaggctgaaa atacactatc taagctcctc 1140
tccaccaaac tgccatattg ccgtgagaat gtgtgtctgg cttatggtag tgaatggtca 1200
gtttatgcag tgggctccca agctcatgtc tccttcttgg atccacggca gccatcatac 1260
aacgtcaagt ctgtctgttc cagggagcga ggcagtggaa tccggtcagt gagtttctac 1320
gagcacatca tcactgtggg aacagggcag ggctccctgc tgttctatga catccgagct 1380
cagagatttc tggaagagag gctctcagct tgttatgggt ccaagcccag actagcaggg 1440
gagaatctga aactaaccac tggcaaaggc tggctgaatc atgatgaaac ctggaggaat 1500
tacttttcag acattgactt cttccccaat gctgtttaca cccactgcta cgactcgtct 1560
ggaacgaaac tctttgtggc aggaggtccc ctcccttcag ggctccatgg aaactatgct 1620
gggctctgga gttaatgaca actccccaaa tgcagagatt tacactaact tccattctca 1680
gtttccttgt ttcttttgat ttttttttc ctaattgtgt gaggctcttg tgttttagtg 1740
ggaacaccaa agtttgccta tagtttaggc acttaatagg aagaagctct gtacagaaat 1800
ctgaaagttg ttttgctttt tgttttcccc tttggtaatc aaaattttac tatcttttat 1860
tatttctggc ttttcaacca aacattgttg ctaatcccta tttttcttta agtgacacac 1920
attctcctgt ctctggcttc ttcaggctga aatgacatag tctttctcac ccttacttca 1980
ctcttgagag gtagggctcc tttataatta catggttgct ctcagacttt ctgtgaaagt 2040
ttgggagctg tgtgtgtctg tgtgtgtgtg agagagagat cttgtctgcg tgtgtgtgtg 2100
tgatcttgtg tgcctgtagg tactgtgtgt cactgaaatt acctggagtg aggattactt 2160
gtaattaaaa tatttataaa agaaacaact ttattcacag agtccagctt tgggactagt 2220
ctgtatcttg ttttttaagt ctaacaacac tgataatagg aagtaaaaac agaaaggaaa 2280
agaaattacc actgggaaaa tctttttagt tagattgtag gcttcctggg gcctcccatg 2340
ccaggactgc aaagtgatcc agccctacct gtcttcccac ctgtgtgtcc cccgtgtggg 2400
aagttggtgt cacttcccct tcccaccctc acatctgctt agccagtagc cacacccccta 2460
aaacatcaga ctcaccatcc aggtgcagct ccagaggcta caaaaggctt catgggactt 2520
gaatccccat cctagcttct ctctccttcc cctcaagacc tgatctggtt ttaaggggcc 2580
tggagctggg agtctcaagt ctgctaagat tcacatccat agccccccgtg gctttgagga 2640
gaatcctctc tgccattctt ccaatctccc cagtgggttt tgctattatt ttctaaattg 2700
ggttaagtct aagaaggtgg gggtgagcag ggggtttatc tgtgtgtagt gagtgcttca 2760
tgtgtggaat attcattttc ttactgcagt gggacttggg gttgaagcca cccctcctac 2820
tctgttggct tagccctgag atggtgacag gctggcctgc agtcagcatc attgtgcatg 2880
tgacagcatc aatgtgatta gtaatttgtc tgttcctccc ttgaactgtc tgtttagtct 2940
gaggttttta aacttgcagg cagctgactg tgatgtccac ttgttccctg atttttacac 3000
atcatgtcaa agataacagc tgttcccacc caccagttcc tctaagcaca tactctgctt 3060
ttctgtcaac atcccatttt ggggaaagga aaagtcatat ttattcctgc accccagttt 3120
tttaacttgt tctcccagtt gtccccctct tctctgggtg taagaaggga aattggaaaa 3180
aaaattatat atatattctc cttttaatgg tggggggcta ctggagagga gagacagcaa 3240
gtccacccta acttgttaca cagcacatac cacaggttct ggaattctca tcttcgaacc 3300
tagagaaata ggtgctataa acagggaatt aagcaaaatg ctggatgcta tagatctttt 3360
aattgtctta atttttttc tattattaaa ctacaggctg tagatttctt agttctcaca 3420
gaacttctat cattttaaac tgacttgtat atttaaaaaa aaaatcttca gtaggatgtt 3480
ttgtactatt gctagaccct cttctgtaat gggtaatgcg tttgattgtt tgagattttc 3540
tgtttttaaa aatgtagcac ttgacttttt gccaaggaaa aaaataaaaa ttattccagt 3600
gcaaaaaa 3608

<210> 100
<211> 1311
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2590967CB1

<400> 100
ggcaggatga acgctgcttt ccaagatggc gacggaggga ggagggaagg agatgaacga 60
gattaagacc caattcacca cccgggaagg tctgtacaag ctgctgccgc actcggagta 120
cagccggccc aaccgggtgc ccttcaactc gcagggatcc aaccctgtcc gcgtctcctt 180
cgtaaacctc aacgaccagt ctggcaacgg cgaccgcctc tgcttcaatg tgggccggga 240

gctgtacttc tatatctaca agggggtccg caaggctgct gacttgagta aaccaataga 300
taaaaggata tacaaaggaa cacagcctac ttgtcatgac ttcaaccacc taacagccac 360
agcagaaagt gtctctctcc tagtgggctt ttccgcaggc caagtccagc ttatagaccc 420
aatcaaaaaa gaaactagca aactttttaa tgaggaaggc tcattgtcat ccccaagcca 480
ggccagttct ccaggtggaa ctgtagtgta gcgacctcac tgctgcgcgc acagtctccc 540
gggacttgga ctcgagggag tgacgaggag gagctccgag ctgcgcctga gccgtgccag 600
ccggcggacc tcaggcggtg gacgtcggcg atagccgtgt ggacggtgac cggctcactc 660
tgcggcgccg tgctcccgct gctcacccaa agaagttgtt tccattttaa accggtcttt 720
tggggctgca gtaaaaaata agaaatggag ttttcttgct ttttactcta aaattcaatg 780
taattaaatt tcatatatat ataatatata catatataca tagtgtaaaa taaaatgttt 840
cttggacaag aaatcccctg aaattcagct gttatagtgc ttcactgttt ttgcactgat 900
ttttctatac cttaggtggt cagaagacaa ccttgaatgc actcatagag aaaactgtta 960
ctttctgacg taatgtaatt caggaagaca gacgctgcaa tcacagattt taaaaaattg 1020
tttgcactta aaaatagttg aatgctggtg gaaagttact ttgcagatgg gtgtaaggac 1080
tcatggccct ctgaggtgcg gcgtgaagat gcccttttta cccgttgacg tttattttac 1140
gtaaaataaa ctgttgtttc caatgcaatc aactctgtat tatatgtata aatattgtaa 1200
ttctgcaatt ggggaaaata gttacttcac tagtaatttt catcatttaa gagtgatatt 1260
tctaattcac aaaagttaat attaaaacta tcttgaatat aaaaaaaaaa a 1311

<210> 101
<211> 2839
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2824491CB1

<400> 101
ggcctgcggg aagccaagat ggcgcatagg ggttctccag gctgcagttg gcgccttatc 60
agtatctaag cggagtgttt tggaaggagt taaggggctg tggcaaacgc cctctccgcc 120
gtcatggccc ggcatcggaa tgttcgaggc tataactacg atgaagattt tgaagatgat 180
gatctctacg gccagtctgt agaggatgat tattgtattt cgccgtcaac agctgctcag 240
tttatttatt cacggcgtga caaacccttcc gttgagcctg tggaagaata tgattatgaa 300
gatctgaaag aatcttccaa ttctgtttca aaccatcagc tcagtggatt tgatcaagct 360
cgtctttatt catgccttga tcacatgaga gaggtacttg gagatgctgt gccagatgaa 420
atattaattg aagcagttct gaagaacaag tttgatgtgc agaaggcttt gtcaggggtt 480
ctggaacaag atagagtgca gagtttgaag gacaagaatg aggcaacagt atctacagga 540
aagatagcaa aaggaaaacc agtagattcc cagacatcgc gaagtgaatc tgaaattgtg 600
ccaaaagttg ctaaaatgac tgtatctgga aagaagcaaa ctatgggatt tgaagtgcct 660
ggagtatctt ctgaagaaaa tggtcatagt ttccacacac ctcaaaaagg accgcccatt 720
gaagatgcca ttgcttcttc cgatgttctt gagactgctt ctaaatctgc taatccaccc 780
cacacgattc aagcatcaga agagcagagt tcaaccccag caccggtgaa aaagtctggc 840
aagctgaggc agcaaataga tgtgaaggcg gaactggaga agcggcaagg agggaagcag 900
ctactcaact tagtggtcat tggtcatgtt gatgctggga aaagtactct gatgggccat 960
atgctttatc ttctgggtaa tataaacaaa agaactatgc ataagtatga acaggagtct 1020
aaaaaggctg gcaaagcttc gtttgcatat gcatgggtct tggatgaaac tggcgaagaa 1080
agggaaaggg gagtaaccat ggatgttggt atgacaaagt ttgaaaccac aaccaaagtt 1140
attacattaa tggatgctcc aggccataag gacttcattc caaatatgat tacaggagca 1200
gcccaggcgg atgtagctgt tttagttgta gatgccagca ggggagagtt tgaagctgga 1260
tttgagactg gaggacaaac acgagagcat ggactcttgg tccgttctct gggagtgacg 1320
cagcttgcag ttgcagttaa taaaatggat caggttaatt ggcaacaaga aaggtttcaa 1380
gagattactg gaaaacttgg gcactttctt aagcaagcag gttttaagga gagtgatgta 1440
ggttttattc ctacaagtgg tctcagtggt gaaaatctaa tcacaagatc tcagtcaagt 1500
gaactcacaa aatggtataa aggactatgt ttattagaac aaattgattc ctttaagcct 1560
ccccagcgat ctattgacaa acctttttaga ttatgtgtgt ccgatgtttt caaagatcaa 1620
ggatctggat tttgcataac tggtaaaata gaagctggtt atatccaaac tggtgaccga 1680
ctactggcaa tgcctcctaa tgaaacttgt accgtgaaag gaatcactct gcatgatgaa 1740
cctgtcgact gggcggcagc aggcgatcat gttagtctta ctttggttgg gatggatatc 1800
atcaaaatca atgttggctg catattttgt ggccccaaag tacccattaa agcttgcact 1860
cgtttcagag cccgaatcct catctttaat attgaaattc ctatcactaa aggatttcct 1920
gtgctgttac actaccaaac tgtcagtgaa cccgccgtta ttaaacgatt gattagtgtc 1980
ttaaacaaaa gcacgggtga agtcacaaag aaaaagccta agtttttgac taaaggccag 2040

aatgcattgg tagagctaca gacacaaaga ccaatagctc ttgagctata taaagacttt 2100
aaagagctgg ggaggttcat gctacgttac ggtggttcta caatagctgc tggtgttgtc 2160
actgagataa aagaatgatg ggtcagaatt tctaccacgt ttctggatac agtgaaatag 2220
ctaacctctg tttcaagaat gcagttatta agtcaaagga acaatgtgca attgatatgt 2280
ttttagatga gagagaaaaa ttaaagctaa aattagctgc aaagaagtat taataatcac 2340
ctctgcaaaa attctaagtt gccaactggc aaagaaagtc taatgttaaa aacaactttg 2400
cctttgaaac gttaataaat ggatttactt tgctaagatt tatggcaagt gtcaaaaata 2460
gtatctgaag atactgaatc atcatgaaat gaactctact tctggccaaa gcacaatgta 2520
tttgcagttt tctcttttga ttcaattata ctgcacatgt tttaaggaaa agtaacttaa 2580
ttgggttttt caggcagttg atatttgacc taagcttttt tttttttttt tttttttttt 2640
tccagttaat gctaagaaaa gatttgggga aggttataat aaaagtattt tgtggtgacc 2700
ataagaatgt ccctccccaa acaagtaaac ttgtgaaagt ttaatttgga attagtggaa 2760
gctgttcctt tgaaagccaa gatattattt aagttgtaaa gccagctaat aaaatgcctt 2820
agtttgagca taaaaaaaa 2839

<210> 102
<211> 1676
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2825460CB1

<400> 102
gggaggcgga ggttgaggtg agaggagatc gcgccattgc actccagcct gggcaacaag 60
ggcgaagttc tctcaaaaga aaaaaaaaaa tcgtcggttt ccttttccca tctttctttc 120
gtgacttata ttccagaacg aagccccaca catcattaat gatttgaatc gtttctaaag 180
tgtttcttaa atcgtttctt aaatcgtttg ttgtttcttg tctaacagtc cagaacacat 240
attacataat ggagccggga gacagactag ggctggcttg atccggccac gcagtccagg 300
aaaggtgctt ttcacccccca agtgcaaaat gatcaatgta ttcttccgat ctacataaac 360
aagcacctcc tggtttcatt ttcgtaaagc aaaacaagca tggaagcttt actgtttcgg 420
ctcttcaaac ttccagcaac tacactgcgc tgcatcggac tcgacgcccg ctggtgacgc 480
acacgctgcg ccggaagtgt gaactgtctg cctccaggct ttgtcatggc ggctgctgct 540
gcacgctgga accatgtgtg ggtcggcacc gagactggga tcttgaaagg ggtaaatctt 600
cagcgaaaac aggcggcgaa cttcacggcc ggaggacagc cgcggcgcga ggaggcagtg 660
agcgccctgt gttggggcac cggcggcgag acccagatgc tggtgggctg cgcggacagg 720
acggtgaagc acttcagcac cgaggatggc atattccagg gtcagagaca ctgcccgggc 780
ggggagggca tgttccgtgg cctcgcccag gccgacggca ccctcatcac atgtgtggat 840
tctgggattc tcagagtctg gcatgacaag gacaaggaca catcctctga cccactcctg 900
gaactgagag tgggccctgg ggtgtgtagg atgcgccaag acccagcaca cccccatgtg 960
gttgccacag gtgggaaaga gaatgctttg aagatatggg acctgcaggg ctctgaggaa 1020
cctgtgttca gggccaagaa cgtgcggaat gactggctgg acttgcgggt tcccatctgg 1080
gaccaggaca tacagtttct cccaggatca cagaagcttg tcacctgcac agggtaccac 1140
caggtccgtg tttatgatcc agcatccccc cagcgccggc cagtcctaga gaccacctat 1200
ggagagtacc cactaacagc catgaccctc actccgggag gcaactcagt gattgtggga 1260
aacactcatg ggcagctggc agaaattgac cttcggcaag ggcgtctact gggctgtctg 1320
aaggggctgg caggcagtgt gcgtgggttg cagtgccacc cttcaaagcc tctactagcc 1380
tcctgtggct tggacagagt cttgaggata cacaggatcc agaatccacg ggtctggag 1440
cataaggatg agccccaaga gcctcaagaa cccaacaagg tgcccctaga agacacagag 1500
acagatgaac tttgggcatc cttggaggca gctgccaagc ggaagctctc gggtttggag 1560
cagccccaag gagctctcca aacgagacgg agaaagaaga agcggcctgg gtccaccagc 1620
ccctgacgcc cctgtgccca ctttgtaaat aaactgctga acacccaaaa aaaaaa 1676

<210> 103
<211> 3206
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2871116CB1

<400> 103
ccagagcgtg cgttcggtgg cccatagggg aagatggcgg ctgctccttt ggaggagcgg 60
gattgagagg atcggggtgg ggagaccaaa caagagagac atttctggct ctgaaggcga 120
acgcttcgct ggccatttag gagctctgct caaagccaga cgtatcctag aaggaaaaca 180
tcaccatggc tacagaaatt ggttctcctc ctcgtttttt ccatatgcca aggttccagc 240
accaggcacc tcgacagctg ttttataagc gacctgattt tgcacaacag caagcaatgc 300
aacagcttac ttttgatgga aaacgaatga gaaaagctgt gaaccgaaaa accatagact 360
acaatccatc tgtaattaag tatttggaga acagaatatg gcaaagagac cagagagata 420
tgcgggcaat tcagcctgat gcaggttatt acaatgatct ggtcccacct ataggaatgt 480
tgaataatcc tatgaatgca gtaacaacaa aatttgttcg gacatcaaca aataaagtaa 540
agtgtcctgt atttgttgtt aggctgcagg aagagtttga aagcctcagt gtccttaaat 600
cgtggactcc agaaggaaga cgcttggtca ctggagcttc tagtggggag tttaccctgt 660
ggaatggact cactttcaat tttgaaacaa tattacaggc tcacgacagc ccagtgaggg 720
ccatgacgtg gtcacataat gacatgtgga tgttgacagc agaccacgga ggatatgtga 780
aatattggca gtcgaacatg aacaacgtca agatgttcca ggcacataag gaggcgatta 840
gagaggccag gtttatacac aatataccat tttctgtagt ccctattgtc atggttaaat 900
tattctctaa gtgtattctg ggtgcagaga tgcatgggct ctgtcagttt ctgggaaact 960
ttctgcaccc tataaacaca atatttttct ttgttttcac acattcacca ttttgctggc 1020
acctttctga agtagtgttg tcccggtatc agcctttgca atatgttaga gatgtactgt 1080
ctgccgcatt ttgcactggt ttttctttt catttatgat taataatgtg tatacgttat 1140
tcctttttat tatctactgt gtaagacaag aatatttcat tccaaataaa gaattcagtc 1200
tttaattatg caactgaata aaatctaaag cctacagaaa acaacttcag aattcacaca 1260
aagtggaaaa aggcttaagt gaagacctgg ttggcttggt tatgccacga cttccaaagg 1320
aaagtatagg actaaaaccc tcacagataa ctggatgtgg caaacattaa cggagtaatg 1380
aatgggttct tcaagctttg cagctgtaag cagatcattg tcaagaagac tctaggactt 1440
ttcttctgat tcactgttga taacatcact tatgcaaatg tatacaataa gtggagttta 1500
aaatattttc agtgagttgt atatttttac acatcagtga ggtatgtata gtaaaactgg 1560
gggaaaaagt tccaaataca agcctgaaga attgctgcag cctcagaata aagctaagca 1620
gcattcttta aggttgtgcc acccatgtgt gggaggaggt tgacatcttt atggaaacat 1680
catccactgt agtcatttgt tcatactttc agaatcttaa cagaaattgt tggatgaaca 1740
tgcttctgct ttgtagattt tgccttagtg tcatgcccat acattgagtt tacacagctg 1800
gtccttcata ggattccaaa gttcaaggga gtttttagag ttagttgaga aacttgatga 1860
tctttcactg ctgggaaaaa ctgactcctt cttgcagcag attctttggc tttacacaca 1920
agtctgaatg tccttatttt aaagtttcc tcaaaggtgc aacattcatg gaatagcttg 1980
ccaggaagat gtgaaacttt tctacagacc tttgaaatgg atgagaaaca ttgtatgtag 2040
ggatgtttag caatcagtct tttaatagac agcccacatt gtttcagctt atttcatgaa 2100
gtgtctgagg cagaagctga tgataatttt gggagcagta ttcgtgtgtg atttaaaaga 2160
ctgcaggaat actgcaaaaa tagaatccat ttattttcac cacttaaggc agcttcatgt 2220
gatttcctcg tatcatagaa aatagagaag gaacatggat agcattagca ctaataatac 2280
acacttgaag ttctcagaat actgatgatt gaaaactcaa acaactgctc tgttgaagtc 2340
ttcttttgat gagatgccta tgttagctga cgacattcac tttaagggct tcttcactgg 2400
attcttccct ctcctgttta taatgcagca cagtgttttt atttttccct gtctgagaag 2460
cacagattat ctgttaaatg ctgacttctt tcccctgctg tgtgtcttca tgtaacagtt 2520
tctcacccac ggataataaa tttgctacat gctctgatga cggcactgtt agaatctggg 2580
actttcttcg ttgccatgag gaaagaattc tccgaggtac gtgtactaac agtactgatt 2640
ggaatattta aatagggaag acatttgtgg ttaaatcatc acaaaaccac aatactggct 2700
tacacctcca ttcaattttt tttacatata cacaccgtct caggctcttc aaaaaaaccc 2760
agcactttct ctgactcaca gtcattttgt aggtttttac taccagtgtt atctttgaat 2820
ttttcagctg taaattaaat acaagagtgc ctcccccttа cttgcttatc tgtatgcatc 2880
ttttagggct gtattccttt tccttccttg tagccagggt acttgttccc aacatattga 2940
cactgtggtt tgatttagat agccgtcatt ctcctggcag tccttttaca atatgaatta 3000
accgacaaga tagaggtatc aaagctacac ttcttagtgt tactattttt gaaagcagtt 3060
ggtttttcag tacaccacat ttgtactaca tggccggctt gttactaagt tcgggtggca 3120
ttgctgcttg tttacttttg ttgattttat aattaataaa cctctatgaa attacttcat 3180
tccgtaactg aaaaaaaaaa aaaaaa 3206

<210> 104
<211> 921
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature

<223> Incyte ID No: 2942212CB1

<400> 104
ggtgctgatg ctgctgccat ttcatcacct ttgcgagcgc acatccatcc ctccgctctc 60
ccggcgcctg ggcctaccca gcttcgggct cccaggccag cgatgcgctc gcggctgagc 120
tagatcctgc cgagccgcgc tctctgaggc gtcggcgggg cgccccctcc cgccgtcccc 180
ggtccgggcc aaggagacct gcagagccgc ggccatggag gccatctggc tgtaccagtt 240
ccggctcatt gtcatcgggg attccacagt gggcaagtcc tgcctgatcc gccgcttcac 300
cgagggtcgc tttgcccagg tttctgaccc caccgtgggg gtggattttt tctcccgctt 360
ggtggagatc gagccaggaa aacgcatcaa gctccagatc tgggataccg cgggtcaaga 420
gaggttcaga tccatcactc gcgcctacta caggaactca gtaggtggtc ttctcttatt 480
tgccattacc aaccgcaggt ccttccagaa tgtccatgag tggttagaag agaccaaagt 540
acacgttcag ccctaccaaa ttgtatttgt tctggtgggt cacaagtgtg acctggatac 600
acagaggcaa gtgactcgcc acgaggccga gaaactggct gctgcatacg gcatgaagta 660
cattgaaacg tcagcccgag atgccattaa tgtggagaaa gccttcacag acctgacaag 720
agacatatat gagctggtta aaggggggga gattacaatc caggagggct gggaaggggt 780
gaagagtgga tttgtaccaa atgtggttca ctcttcagaa gaggttgtca aatcagagag 840
gagatgtttg tgctagtcag ttctttttatt tccaaaacat gctctcctac ttgaactgaa 900
aagtaagaga aataaataga a 921

<210> 105
<211> 1367
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3685151CB1

<400> 105
aagaggcacg tgcgctgctg aatggagctg gtcgctggtt gctacgagca ggtcctcttt 60
gggttcgctg tacacccgga gcccgaggct tgcggcgacc acgagcagca atggactctt 120
gtggctgact tcactcacca tgctcacact gcctccttgt cagcagtagc tgtaaatagt 180
cgttttgtgg tcactgggag caaagatgaa acaattcaca tttatgacat gaaaaagaag 240
attgagcatg gggctctagt gcatcacagt ggtacaataa cttgcctgac attctatggc 300
aacaggcatt taatcagtgg agcggaagat ggactcatct gtatctggga tgcaaagaaa 360
tgggaatccc tgacgtcaat taaagctcac aaaggacagg tgaccttcct ttctattcac 420
ccatctggca agttggccct gtcggttggt acagataaaa ctttaagaac gtggaatctt 480
gtagaaggaa gatcagcatt cataaaaaat ataaaacaaa atgctcacat agtagaatgg 540
tccccaagag gagagcagta tgtagttatc atacagaata aaatagacat ctatcagctt 600
gacactgcat ccattagtgg caccatcaca aatgaaaaga gaatttcctc tgttaaattt 660
ctttcagagt ctgtccttgc agtggctgga gatgaagaag ttataaggtt ttttgactgt 720
gattcactag tgtgcctctg cgaatttaaa gctcatgaaa acagggtaaa ggacatgttc 780
agttttgaaa ttccagagca tcatgttatt gtttcagcat cgagtgatgg tttcatcaaa 840
atgtggaagc ttaagcagga taagaaagtt cccccatctt tactctgtga aataaacact 900
aatgccaggc tgacgtgtct tggagtgtgg ctagacaaag tggcagacat gaaagaaagc 960
cttcctccag ctgcagagcc ttctcctgta agtaaagaac agtccaaaat tggcaaaaag 1020
gagcctggtg acacagtgca caaagaagaa aagcggtcaa aacctaacac aaagaaacgc 1080
ggtttaacag gtgacagtaa gaaagcaaca aaagaaagtg gcctgatatc aaccaagaag 1140
aggaaaatgg tagaaatgtt ggaaaagaag aggaaaaaga agaaaataaa aacaatgcag 1200
tgaatcacag atgtctcctg aaagaactct tttagatgaa atcattctac tcaaatgtac 1260
cttaattttt tttttttccc tgagtaaaag caagaaattt cttcctttgg aaaaaatata 1320
tatattaaaa aaccactttt agatggtttt ttttaaaaaa aaaaaaa 1367

<210> 106
<211> 1560
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4881515CB1

<400> 106
aggcggactg gggaggcggc ggcctggctc ggcctggcct ggcctgtcag ggcgcgggcg 60
gcggcggctc cagcaccatg tccctgcagt acggggcgga ggagacgccc ctcgccggca 120
gttacggcgc ggccgattcg tttccaaagg acttcggcta cggcgtggag gaggaggaag 180
aggaggcggc ggcggcgggc ggaggggttg gggcaggggc aggcggtggc tgtggtccgg 240
ggggcgctga cagctccaag ccgaggattc tgctcatggg actccggcgc agcggcaagt 300
cctccatcca gaaggtggtg tttcataaga tgtcacccaa cgagaccctc ttttggaaa 360
gtaccaacaa gatttataag gatgacattt ccaatagctc ctttgtgaat ttccagatat 420
gggatttcc tgggcaaatg gactttttg acccaacctt tgactatgag atgatcttca 480
ggggaacagg agcattgata tacgtcattg acgcacagga tgactacatg gaggctttaa 540
caagacttca cattactgtt tctaaagcct acaaagttaa cccagacatg aattttgagg 600
tttttattca caaagttgat ggtctgtctg atgatcacaa aatagaaaca cagaggaca 660
ttcatcaaag ggccaatgat gaccttgcag atgctgggct agaaaaactc catcttagct 720
tttatctgac tagtatctat gaccattcaa tatttgaagc ctttagtaag gtggtgcaga 780
aactcattcc acaactgccg accttggaaa acctattaaa tatctttata tcaaattcag 840
gtattgaaaa agcttttctc tttgatgttg tcagcaaaat ctacattgca acagacagtt 900
cccctgtgga tatgcaatct tatgaacttt gctgtgacat gatcgatgtt gtaattgatg 960
tgtcttgtat atatgggtta aaggaagatg gaagtggaag tgcttatgac aaagaatcta 1020
tggcaattat caagctgaat aatacaactg tcctttattt aaaggaggtg actaaatttt 1080
tggcactggt ctgcattcta agggaagaaa gctttgaaag aaaaggttta atagactaca 1140
acttccactg tttccgaaaa gctattcatg aggtttttga ggtgggtgtg acttctcaca 1200
ggagctgtgg tcaccagact agtgcctcca gtctgaaagc gctgacacac aatggcacgc 1260
cacgaaacgc catctagtct gaatcccagc gtcggggctc tgtgccagct tactcttcac 1320
tccagggtcg gatgccacgt gctacaggac atgggagctg ctgcttgtgg gaatctggtg 1380
cctgttccac tagagacaag ggtagagtt tctcatttgg atgaaaaccc cttcaactgg 1440
tggtgtacaa ctgaagctac tatatctttt ttgaaaatgg caaaaaaaaa aaaaaaaaat 1500
tctggagacc acagaactca agtgtgtgtt tctcctcttt tgggtcccct ttaagtagtt 1560

<210> 107
<211> 1495
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5324681CB1

<400> 107
gaagaggctc tgggctggca catgtgtatg gcggtgaggc gggcgggtac atggcgggct 60
ctgtgggact ggcgttgtgc gggcagacgt tggtggtgcg gggcggcagc cgattcctgg 120
ccacctccat agcaagcagt gatgatgaca gcctcttcat ctatgactgc agtgctgcag 180
aaaagaagtc acaagaaaat aaaggggagg acgcgccctt ggaccagggg agcggtgcga 240
ttctggcgtc caccttctcc aagtctggca gctattttgc tttaaccgat gacagtaagc 300
gtctgattct tttccgtaca aaaccatggc aatgtctgag tgtcaggacc gtggcaagga 360
ggtgtacagc cctgactttc atagcctcgg aggagaaggt cttggtggcc gacaagtctg 420
gagacgtcta ctccttttcg gtgctggagc cacacgggtg tggccgtcta gagctggggc 480
acctgtctat gctgttagat gtggctgtga gtcctgatga ccgcttcatc ctcactgccg 540
accgggacga gaagatccga gtcagctggg ccgcggcgcc ccatagcatc gagtccttct 600
gcttggggca cacagagttt gtgagccgta tctccgtggt gccaactcag cccgggctgc 660
ttctgtcctc ctctggggac ggcaccctga ggctctggga gtacaggagc ggccgccagc 720
tgcactgctg tcacctggcc agtctgcagg agctggtgga cccccaggcc ccccagaagt 780
ttgccgcgtc caggattgca ttctggtgcc aggagaactg cgtggcgctc ctgtgcgacg 840
gcactcctgt ggtctacatc ttccagctgg acgcccgcag acagcagttg gtgtacaggc 900
agcagctggc gttccagcac caagtgtggg acgtggcttt cgaggagacc caggggctgt 960
gggtgctcca ggactgccag gaagcccccc tggtgctcta caggcctgtg ggcgaccagt 1020
ggcagtctgt tcctgagagc accgtgttaa agaaagtctc tggtgttctt cgtgggaact 1080
gggccatgct ggaaggctct gccggcgcag acgccagctt cagcagtctc tacaaggcca 1140
cgttcgacaa cgtgacctcc tacctgaaga agaaagagga gagactgcag cagcagctag 1200
agaagaagca gcggcgccgg agtcccccgc ctgggcccga cggcatgcc aagaagatga 1260
gaccggggga ggcgacgcta agttgctgat cgtggcggtc tgtttctgtc gactgtggac 1320
cacttatgtg cgatccgtgg accacttgcg tgcgatctgt cggccgacga tgagcttgtt 1380
cggatgtagc tccatcgtaa gtcgaggagc atctgtgatt tgtcctctgc ttatgggata 1440
tgtttttccg ctactgagtc tgtgtagtaa atttttgact aggaaaaaaa aaaaa 1495

<210> 108
<211> 1919
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5387651CB1

<400> 108
cgccctgcat gcgagttggg ccgcgggcgg ggttggagcc tactcggggc gactgcgatg 60
gacgccttag aaggagagag ctttgcgctg tctttctcct ccgcctctga tgcagaattt 120
gatgctgtgg ttggatattt agaggacatt atcatggatg acgagttcca gttattacag 180
agaaatttca tggacaagta ctacctggag tttgaagaca cagaagagaa taaactcatc 240
tacacaccta tttttaatga atacatttct ttggtagaaa aatacattga agaacagctg 300
ctgcagcgga ttcctgagtt caacatggca gccttcacca caacattaca gcaccataag 360
gatgaagtgg ctggtgacat attcgacatg ctgctcacct tcacagattt tctggctttt 420
aaagaaatgt ttttggacta cagagcagaa aaagaaggcc gaggactgga cttaagcagt 480
ggcttagtgg tgacttcatt gtgcaaatca tcttctctgc cagcttccca gaacaatctg 540
cggcactagg tcctacctcc agccaatgaa tggatcatt ctggatgtca ccagcccaat 600
aggctcagct catgatgaca gaacacatct tggaaagact gactctgtta tgtaactctt 660
catttatgtt aagtattaat aggtcaaaac caaaatgacc taaccctcct ggacctattt 720
atcctgaaac accttcttgt attcattaac catagtactc ctccccacct caagtagaca 780
cctctctcag gagcttctga gtcagacgcc tctggagcga gccctatgtc aggcactcca 840
cctggggggc ccttccccag catacctgct ggtgtgtaag tgtggactaa cccgccgcca 900
ccaccctctg ttccagcagg ctctgcatga atctttgtgc acttgcacct ctttttcaca 960
tgggccacag tttcagtact tcagcctcag tggggttcct gatgtttatc tagggtgtta 1020
ctcaagccca gtttgagatt ttggagtctc ctgtgatcac atcttgtctc ggctgtagga 1080
atcaacagaa ggagacgtcc tctacataaa agctccatgt gaaaagctac tcctagtctt 1140
aacatttgca gtccttgtgt cactgtcttc tggtcctgat gtagtcccac tgtttctaga 1200
agtctctttt aagcattatt tttgaaaaaa aaaatatttt tatagatgaa tactcaggct 1260
aacctagtgg atgtgatctt ggaacttcca tgattatcca cttaaagatc aaagtattat 1320
atgctgtgtg ctttttaggt gtttgttagt actgtgaagg caaaaatgct ttctacattg 1380
acattcattc ctattttact gggcacctat gaatgtatgc tgtgtgctag aaatagacta 1440
aaacatattc ctatagcatg ttagtgtgtt tgcatgtttg ctgaaaatcc tttgtgtata 1500
aaccagtttg taaggttctc tgggttaggt agggactctg cagtttcttc ctgtcaaaat 1560
ctctcctacc aagatggtgt tccactgtcc agcccagcat gagtagcagg tagagcacag 1620
ctttactggc tgtttgtatg ctttggttta gtgcaatgtg tggtagatta cttatcagaa 1680
aacatatatg tcatctctag aacgaagaaa aagcatagta gttcaattcc cagtgtgtcc 1740
ctttgatttt ttttttttaa tagtaaaaat aagaatctgt actgactttt cacttggcca 1800
ttctggtttt aaaggacaag ctacaagctc tgtgtttctg tactgatgtg tcacttatta 1860
aatacttttg taccatgagt aaaacttcag gtgtttcgca agaaccacca ttctcaaaa 1919

<210> 109
<211> 2941
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5595679CB1

<400> 109
attaggctaa taggacagca cttgaatggc ttagggctca accagactgt tgatctcctc 60
atgcaagagt caggatgtcg tttagaacat ccttctgcta ccaaattccg aaatcatgtc 120
atggaaggag actgggataa ggcagaaaat gacctgaatg aactaaagcc tttagtgcat 180
tctcctcatg ctattgtggt aagaggcgca cttgaaatct ctcaaacgtt gttgggaata 240
attgtgagga tgaagttttt gctgctgcag cagaagtacc tagaatacct ggaggatggc 300
aaggtcctgg aggcacttca agttctacgc tgtgaattga cgccgctgaa atacaataca 360
gagcgcattc atgttcttag tgggtatctg atgtgtagcc atgcagaaga cctacgtgca 420
aaagcagaat gggaaggcaa agggacagct tcccgatcta aactattgga taaacttcag 480
acctatttac caccatcagt gatgcttccc ccacggcgtt tacagactct cctgcggcag 540
gcggtggaac tacaaaggga tcggtgccta tatcacaata ccaaacttga taataatcta 600

```
gattctgtgt ctctgcttat agaccatgtt tgtagtagga ggcagttccc atgttatacg 660
cagcagatac ttacggagca ttgtaatgaa gtgtggttct gtaaattctc taatgatggc 720
actaaactag caacaggatc aaaagataca acagttatca tatggcaagt tgatccggat 780
acacacctgc taaaactgct taaaacatta gaaggacatg cttatggcgt ttcttatatt 840
gcatggagtc cagatgacaa ctatcttgtt gcttgtggcc cagatgactg ctctgagctt 900
tggctttgga atgtacaaac aggagaacta aggacaaaaa tgagccagtc tcatgaagac 960
agtttgacaa gtgtggcttg gaatccagat gggaagcgct ttgtgactgg aggtcagcgt 1020
gggcagttct atcagtgtga cttagatggt aatctccttg actcctggga aggggtaaga 1080
gtgcaatgcc tttggtgctt gagtgatgga aagactgttc tggcatcaga tacacaccag 1140
cgaattcggg gctataactt cgaggacctt acagatagga acatagtaca agaagatcat 1200
cctattatgt cttttactat ttcaaaaaat ggccgattag ctttgttaaa tgtagcaact 1260
cagggagttc atttatggga cttgcaagac agagttttag taagaaagta tcaaggtgtt 1320
acacaagggt tttatacaat tcattcatgt tttggaggcc ataatgaaga cttcatcgct 1380
agtggcagtg aagatcacaa ggtttacatc tggcacaaac gtagtgaact gccaattgcg 1440
gagctgacag ggcacacacg tacagtaaac tgtgtgagct ggaacccaca gattccatcc 1500
atgatggcca gcgcctcaga tgatggcact gttagaatat ggggaccagc accttttata 1560
gaccaccaga atattgaaga ggaatgcagt agcatggata gttgatggtg aatttggagc 1620
agacgacttc tgtttaactt aaaattagtc gtattttaat ggcttgggat ttggtgcaaa 1680
caaacatgat tgatagctgg acagacatgc tcgtcatgaa aaaagaacca tttctgaagc 1740
ccgattgggg ccaaacattt acaccttgct tcatagtaac cagttgagat gaagcacgtc 1800
gttagaacgt tgttggacac catgttgaat tattccccca tcggttgtga agaactgtgc 1860
tacattcagg cttacccatt gaactcagta tatatatttt tttccttcct gtcttttgtc 1920
tggcaggata ccattcttgt tgctcttctg tgtaatgaag tttaaatgct tgtttggaaa 1980
actttattta acagtttaga aggcttgata gaaagagtgc attagtctga agagtataca 2040
ttggatagga aagaatttcc ttcttttgtt tctccaaatc tttccgcctt atttagcttg 2100
agatctttgc agcttggttc atggattcta gccttgcccg ttgcgcagta tatactgatc 2160
cagatgataa accagtgaac tatgtcaaaa gcactctcaa tattacattt gacaaaaagt 2220
tttgtacttt tcacatagct tgttgccccg taaaagggtt aacagcacaa ttttttaaaa 2280
ataaattaag aagtatttat aggattaaag tgacttcatt tgtatacatt tggaatctaa 2340
accagcttaa aaacagtttc ctcaatgact tagatacaca gtataactga tgctcttctg 2400
gaataccaca tgagacatgg tcagaaacag tgcttggaag gacattacac aagaaattca 2460
gagtaatgct ttgaagattt ccccccttttt gttttattcc tgaaggaaca tcagtacccg 2520
atcttgaaga aattcaagat tcaaaaagaa ttttaaatac accaacatga gacatcagta 2580
gtcagttggt tttcagtaaa gcttgttcca agttgttctc aacttaggaa gtaattttgg 2640
tgtgatctag caaaagagta ggaatcagcg atacaaccac tttggaagtt tatagtataa 2700
ttgaaattat tagaagaatt cagcaggtta cagacatact taaactggga ttaaaacctc 2760
atagtcattt ttcttaattg cccttaatat tttgacatat agggatacat aaatttaaag 2820
aatatttttt ctcagttttt tcagatattg ccatactgaa cctcattcta aactggtgct 2880
gtggatagtc tttccctctc ccctcctgtt ttagtttaag gaaaggtttc cttcatggaa 2940
a                                                                 2941
```

<210> 110
<211> 710
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 5782457CB1

<400> 110

```
ctcgcggtcg cctggccgtt gtcattgtcc ctccgctgtc accttttcaa gccccaggct 60
ggctgcttca gaagcccctg accccatgga ggagtgggac gtgccacaga tgaagaaaga 120
ggtggagagc ctcaagtacc agctggcctt ccagcgggag atggcgtcca agaccatccc 180
cgagctgctg aagtggatcg aggacgggat ccccaaggac cccttcctga accccgacct 240
gatgaagaac aacccatggg tggaaaaggg caaatgcacc atcctgtgag ccccgcaccc 300
ggcccctctc acaccatcct gtgagaccac gcccggcccc actcccacca tcttgtaaga 360
ctgtgcccag ccccactcac tccatcctgt gagtcccact cccagcccca ctcccaccat 420
cctgtgagcc catgcccggc cccactcaca ccaacctgtg agccccactc ccggccccac 480
tcacaacatc ttgtaagact gtgcccggcc ccattcactc catcctgtga gaccacgccc 540
ggccccactc actctatcct gtgagaccac gcctggcccc actcccacca tcctgtgagc 600
cccactcctg gccccactca caccatccta tgagcccacg cccggcccca ctcccaccat 660
cctgtgaacc ccactccact cgcacgtgat tacagtctgt aaaggtgtga           710
```

<210> 111
<211> 1351
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 760677CB1

<400> 111
ccaggcctag tggatagaca gggtccaaaa tgtgaccctt ctaggctggt atcaccatgg 60
gggcgtcatg ggctgaggat tctgcagata ggacatcacc acggcagaga tggacagcct 120
gagacaggag agaggtgtca gtctaggatg gccaggctgg gtccccccac cccttactca 180
agagtcactt gttctgtagg gcaagtctca catgaagcta ctcatcattt gcttcgtgtc 240
ctccgagctc cgagagttgg caaagctgat gaaggagcag tagacagcga cccaagcaca 300
ccacttcagc tagggaaaga ccgatttaag gctgcaagga aggagtcctg ggagcatggc 360
tttccctgag ccaaagccgc ggcctccaga gctgccgcag aaacggttga agacgctgga 420
ctgcgggcag ggggcagtgc gagccgtacg atttaatgtg gatggcaatt actgcctgac 480
gtgcggcagt gacaagacgc tgaagctgtg gaacccgctt cgggggacgc tgctgcggac 540
gtacagcggc cacggctacg aggtgctgga tgcggccggc tcctttgaca acagtagtct 600
ctgctccggc ggcggggaca aggcggtggt tctgtggaat gtggcatcag ggcaggtcgt 660
gcgcaaattc cggggccacg cagggaaggt gaacacggtg cagtttagtg aagaggccac 720
agttatcctg tccggctcta ttgattccag tatccgctgt tgggattgcc gctcacggag 780
gcctgagcca gtgcagacgc tggatgaggc cagagatggc gtgtccagtg tgaaggtgtc 840
agaccacgag atcctggcag gctccgtgga tggccgcgtg agacgctatg acctaaggat 900
ggggcagctc ttctcagact acgtgggcag ccccatcacc tgcacctgct tcagccggga 960
tgggcagtgc accctggtgt ccagcctgga ctccacattg cggctcctgg acaaagacac 1020
aggggagctg ctgggcgagt acaagggcca taagaaccag gaatacaagc tggactgctg 1080
cctgagcgag cgtgacacac atgtggtcag ctgttctgag gacgggaagg tgttcttctg 1140
ggacctggtg gagggtgcgc tggctctggc cctgcctgtg ggttccggtg tggtgcagtc 1200
gctggactac cacccaacag agccctgcct gctgaccgcc atgggaggca gcgtccagtg 1260
ctggcgagag gaggcctatg aggcagagga tggagcaggc tgaagccagg ggacccacca 1320
acaggaccaa ggaccgagac acagacatgg c 1351

<210> 112
<211> 1783
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1348567CB1

<400> 112
cccacgcgtc cgcggacgcg tgggctgaag gctgtggcgc gcggctgtcc ccattcccac 60
gtgaagcgct acgctagcat cgctcggctg gcggctccca gctcgccgcg gagcagtccc 120
ggcagcagcg ggggaccgga agtggctcgc ggaggctcag aagctagtcc cggagcccgg 180
cgtgtggcgc ctcggagcgc ggtgacggcg ccatgtccct aatctgctcc atctctaacg 240
aagtgccgga gcacccatgt gtatcccctg tctctaatca tgtttatgag cggcggctca 300
tcgagaagta cattgcggag aatggtaccg accccatcaa caaccagcct ctctccgagg 360
agcagctcat cgacatcaaa gttgctcacc caatccggcc caagcctccc tcagccacca 420
gcatcccggc cattctgaaa gctttgcagg atgagtggga tgcagtcatg ccgcacagct 480
tcactctgcg ccagcagctg cagacaaccc gccaagagct gtcacacgct ctgtaccagc 540
acgatgccgc ctgccgtgtc attgcccgtc tcaccaagga agtcactgct gcccgagaag 600
ctctggctac cctgaaacca caggctggcc tcattgtgcc ccaggctgtg ccaagttccc 660
aaccaagtgt tgtgggtgcg ggtgagccaa tggatttggg tgagctggtg ggaatgaccc 720
cagagattat tcagaagctt caagacaaag ccactgtgct aaccacggag cgcaagaaga 780
gagggaagac tgtgcctgag gagctggtga agccagaaga gctcagcaaa taccggcagg 840
tggcatccca cgtggggttg cacagtgcca gcattcctgg gatcctggcc ctggacctct 900
gcccgtccga caccaacaag atcctcactg gtggggcgga taaaaatgtc gttgtgtttg 960
acaaaagttc tgaacaaatc ctggctaccc tcaaaggcca taccaagaag gtcaccagcg 1020
tggtgtttca cccttcccag gacctggtgt tttctgcttc ccccgatgcc actatcagga 1080
tttggtcggt ccccaatgcc tcttgtgtac aggtggttcg ggcccatgag agtgctgtga 1140

caggcctcag ccttcatgcc actggcgact atctcctgag ctcctccgat gatcagtact 1200
gggctttctc tgacatccag acagggcgtg tgctcaccaa ggtgacagat gagacctccg 1260
gctgctctct cacctgtgca cagttccacc ctgacggact catctttgga acaggaacca 1320
tggactctca gatcaagatc tgggacttga aggaacgtac taatgtggcc aacttccctg 1380
gccactcggg ccccatcact agcatcgcct tctctgagaa tggttactac ctggctacag 1440
cggctgatga ctcctctgtc aagctctggg atctgcgcaa gcttaagaac tttaagactt 1500
tgcagctgga taacaacttt gaggtaaagt cactgatctt tgaccagagt ggtacctacc 1560
tggctcttgg gggcacggat gtccagatct acatctgcaa acaatggacg gagattcttc 1620
actttacaga gcatagcggc ctgaccacag ggtgtggcctt cgggcatcac gccaagttca 1680
tcgcttcaac aggcatggac agaagcctca agttctacag cctgtaggcc ctggcccttc 1740
tgatggaagc tgggcctcat ctcagtagag ggtagaatt agg 1783

<210> 113
<211> 3453
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1751354CB1

<400> 113
ggcttgcgca ctacgtcccc agccagaggc tcctacccgg tcggggactt ccggaacgcc 60
ggggtgtggt tccgggtcgt gtgcggctcg gggtaatagg gctgctgctc ggccggccgg 120
cggcggcgag cagcaggggc atgagggcta acccgggaag cggcagctga gcgggccggg 180
aggagcgccg gtccccgtgg atcccgagag tgcagagctc ggggcagggg ccgggaggcg 240
tgggggagcc gggccctccc ctcaggaacg tgtcccgggg ccgacccggc ccgtagtgtg 300
gaagcagctt caggtaggtg agctcgtgaa acaatatgaa gaggagaaaa tagcctttta 360
aggaaattgg cccacagaaa ggatggcctt cttggacaat ccaactatca ttctagctca 420
tattcgacag tcacatgtga ccagtgatga cacgggaatg tgtgagatgg ttctcattga 480
tcatgatgtt gacctagaga agattcatcc tccttcaatg cctggagaca gtgggtcaga 540
aattcaggga agcaatggtg agactcaggg ctatgtatat gcccagtcag tcgatattac 600
ctcaagttgg gactttggta ttagaagacg ctcaaacaca gctcaaagat tagaacgact 660
ccgaaaagag agacaaaacc agatcaaatg caaaaatatt cagtggaaag aaagaaattc 720
taagcaatca gcccaggagt taaagtcact gtttgaaaaa aatctctca aagagaagcc 780
tccaatttct gggaagcagt cgatattatc tgtacgccta gaacagtgcc ctctgcagct 840
gaataaccct tttaacgagt attccaaatt tgatggcaag ggtcatgtag gtacaacagc 900
aaccaagaag atcgatgtct acctccctct gcactcgagc caggacagac tgctgccaat 960
gaccgtggtg acaatggcca gcgccagggt gcaggacctg atcgggctca tctgctggca 1020
gtatacaagc gaaggacggg agccgaagct caatgacaat gtcagtgcct actgcctgca 1080
tattgctgag gatgatgggg aggtggacac cgatttcccc ccgctggatt ccaatgagcc 1140
cattcataag tttggcttca gtactttggc cctggttgaa aagtactcat ctcctggtct 1200
gacatccaaa gagtcactct ttgttcgaat aaatgctgct catggattct cccttattca 1260
ggtggacaac acaaaggtta ccatgaagga aatcttactg aaggcagtga agcgaagaaa 1320
aggatcccag aaagtttcag gccctcagta ccgcctggag aagcagagcg agcccaatgt 1380
cgccgttgac ctggacagca ctttggagag ccagagcgca tgggagttct gcctggtccg 1440
cgagaacagt tcaagggcag acggggtttt tgaggaggat tcgcaaattg acatagccac 1500
agtacaggat atgcttagca gccaccatta caagtcattc aaagtcagca tgatccacag 1560
actgcgattc acaaccgacg tacagctagg tatctctgga gacaaagtag agatagaccc 1620
tgttacgaat cagaaagcca gcactaagtt ttggattaag cagaaaccca tctcaatcga 1680
ttccgacctg ctctgtgcct gtgaccttgc tgaagagaaa agccccagtc acgcaatatt 1740
taaactcacg tatctaagca atcacgacta taaacacctc tactttgaat cggacgctgc 1800
taccgtcaat gaaattgtgc tcaaggttaa ctacatcctg gaatcgcgag ctagcactgc 1860
ccgggctgac tactttgctc aaaaacaaag aaaactgaac agacgtacga gcttcagctt 1920
ccagaaggag aagaaatccg ggcagcagtg acactggcct ccagcctcaa tctgttccgt 1980
agctcagagc ctgcctgcca gggccaagtg ccctagagcc cacccggtgt cctgaagtcc 2040
tcggggggag gccagcccct ggctcactgg cacagggcag gtgggctctc ggggaaggtg 2100
tcgggggccc cctaggaggg agcgctgggg acattgccat gggacggaag tctgcttggc 2160
agtggctttg ataagcgatg cttgggggtc agaccacccc ctagaggagc cacgtgccgc 2220
ccagccacct tcaatgcctg ccaccctgcc cgaggatgta cagagccgtg cccacacatt 2280
tccttgcaac ttgatcaaat ttcttaaagc aaacaacaaa aatgtacatt tctgtttttc 2340
cttttaataa acaggtgtac tctttatcat ggttggtatg atggaccatt ctttggggcg 2400
gaggattgat tatgttactc tctttaaaat ctgttcccat attgaacagg cagattggaa 2460

aagctatggt tcgatttctc agaagaaatg tttaggtctt agtcaatagt tttaactatg 2520
ccatttgttt aaatgagtgc atttgcttcg agggtagtgt cttactaaaa gttaggaaca 2580
gagacctagt ggtgtgtcca aggccgtgtc actttcccct tcagcacacc ccagcttctg 2640
acctcagagc ccaggagctg cgtggacagt gtggggtgcc aggaggaggg gcggtggctg 2700
gtcctcaggc acgctgcact cccagccaga catggtcttt ccgtttctta agtagcaagt 2760
gtaggtttca gctggcagtt ccacctgcat gttctctgct tcgctgcctt ggaaggggcc 2820
acattcccca ttcctcttct ccttacagcg cctgcctcct ttttcaagca ggcggaaagc 2880
tgctgtttct cacgtttcag ggagaggggt gagcggaggg agacctgtgt ccgtgccgtc 2940
cggctccctg ggtgggaaca ggcaagggat cagatgcccc tgacaccacg cctctggcca 3000
caccagatgc ctctgcagtc ctcgacagcc tcttcagtgt ccctcctgcg gtgatgtcct 3060
tactgtcccc agccagggcc ggggaccggt gtttcactga ggacctgcat tagaaacatt 3120
ttttaaattg ttgtacagga agagatgtgt ctaaaacagc atcttaaagc tgagtgtatt 3180
tctttgcaca aggggtcatg ctgatgaatt cttctttcat tctgatcttt gttcagccaa 3240
caggagcgtc cttttctaat gtcttccatt cctacccccc acccaaaaac aaaagaaata 3300
tttgtagctt gctatctgta tttgaatttt tagcaatttt atatttagat actttgaaaa 3360
atgtaaatga ctaatttggt cattaaatct tgtgacatat tcgatattaa aatgatatta 3420
aaataaaagt catataaata cacaaaaaaa aaa 3453

<210> 114
<211> 2663
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 1976780CB1

<400> 114
gaaaaaggtt cgaaagaact ggttgtcttc ttgggcggtg ttgcagggtt catctttact 60
ttttaccaaa actcaaggaa gtagcacaag ttggtttggc agtaatcagt ccaaaccaga 120
gttcacagtg gacctcaagg gggcaacaat tgagatggct tcaaaggata aatccagcaa 180
aaagaatgta tttgagctga aaactcgtca aggaacagaa ctgctaattc agtctgacaa 240
tgacactgtt attaatgatt ggtttaaagt tcttagtagt acaatcaata atcaggcagt 300
agaaactgat gaaggaattg aagaggagat accggattca ccaggaatag aaaagcatga 360
taaagaaaag gaacaaaagg atcccaaaaa gcttcgttcc tttaaagtat ctagcataga 420
ttcttcagaa cagaaaaaaa ccaagaaaaa cttaaagaag tttcttacac gacgccccac 480
tttgcaagct gttcgtgaaa aaggttatat taaagatcag gtatttggat ccaatctcgc 540
taatctgtgt cagagagaga atggcacagt accaaagttt gtgaagttat gtattgaaca 600
tgttgaagaa catggtttgg atattgatgg gatatacaga gtaagtggca acctcgcagt 660
gatccagaaa ctaaggtttg cagtcaatca tgatgagaaa ttggacttga atgacagtaa 720
atgggaagat attcatgtca ttactggagc cctcaaaatg ttttttcgag aattaccaga 780
acctcttttt acatttaatc attttaatga ttttgttaat gcaattaagc aagaaccaag 840
acagcgagtc gctgctgtta aggacctaat cagacagttg ccaaagccaa accaagacac 900
aatgcagatt cttttccgac atctcagaag agttatagaa aatggagaga aaaatcgaat 960
gacctatcag agtatagcaa ttgtttttgg tcccactcta ttaaaaccag aaaaagagac 1020
tggtaatata gcagttcata ctgtgtacca gaatcagatt gtagaattaa ttcttctgga 1080
actgagttcc atcttcggac gttgattctt actgaagaca acctgtggaa tagaagctgg 1140
attccatcag atttcaaatg tttatacaca atgtatttta ttttttggac caagcagtga 1200
ctctttgatt ttgcactttt tttttgaggg atcagaaggg aaggggagag tcgagatgtg 1260
tgttaggccc tcatatttgc tgctttgttg caagttgata taactgcgtg taattatgaa 1320
ttcattttat cctgaatgtt tgcatttcat actctgaatt tcagtaaaaa tcaaaactta 1380
aaattctaac cagtcatata cactggataa tttggtaaga aaactgtatt ttttttccct 1440
gaaattggat aatgtacttt cttctcaaga ttcatgactt gatagaacaa tactttcagt 1500
tatgttgcaa aggctcttgg gcattttaaa caaaatgaag tatatccatt ttgaaacctg 1560
tgtatttctt tttcggggtt tctgcatgca gtggcagtct taagtgccaa aattcattat 1620
aaccccaaaa taaccccttg atgaaggctt gctgtctttt actgtgttac acagcatcct 1680
tactggatat cttagttgct tgtttgggca gcacactaat attacttaaa acactgtgat 1740
atactggagt tttagttagc ggaagtcagt tcagggcatt ttagggctgt cttgctatac 1800
tgaattgtag ctaacaatcc taattatatc tagtaccata ctgagttatt ggtatgaccc 1860
tgtggaaaca cacattattt tatgtaaata taggctaaag acttaatgtc ctttagcttg 1920
tgtatataat tgtgttgtat agtctcagag tacattctaa ccctacattt ctaatcattg 1980
ttattggtaa tcttttctgt gaatattagg tttcctccag aaatggtccg ttatttggga 2040
aagttaactg tgtgcacttt tagatattaa ctacatttac aggcaaatca ctgtaatgag 2100

aatggtactg gaaaaatact gaatagactt gctaaatggc acatgcacta caagaggaac 2160
cttttgggtt atttaatatg tacagaaaac attagaaaaa atttattaca gaattctaat 2220
tccagtatga atagtggaaa cccatctgta aattagatgg atgttggatg gaaaatgaca 2280
ttgctaaatt tgagaatttc tttttaccta ctaatgtaga ttgctttgta taataaaaca 2340
cagggtttgg aaggttttgt tacagggagc atggtctgtt gaagattttt aaaatgtatt 2400
tttctagatt aacttctgta catgaaatgt ctaataaaac tataagaggt ttagagattt 2460
ttccattgga aatgtgcatt ttggtttcta attttttgt tttttcattt actggcatac 2520
tgttatacct catttttaaa aatcaactga atccaatatt tcctgtggca aataacactt 2580
tcctcatttc atacctttc tcctctcttc catgccaaca tttctccacc cacaacgtac 2640
actgtttatt tctcatcaat att 2663

<210> 115
<211> 1218
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2048234CB1

<400> 115
gcctgttgca gccatggtgc attgcagttg cgtgttgttc agaaagtatg gaaatttcat 60
cgataagcta agactcttca ccaggggagg atccggtgga atgggttatc ctcgtttagg 120
tggagaaggt ggaaaaggtg gtgatgtctg ggttgtagcc cagaacagaa tgactttaaa 180
acaacttaaa gacaggtatc ctcggaaacg gtttgtggct ggagtaggag caaacagcaa 240
aattagtgca ctgaaaggct ccaaaggaaa agactgggaa atccctgtgc ctgtgggtat 300
ttcagtaact gatgaaaatg gtaaaattat aggagaactc agtaaagaaa atgacagaat 360
tttggtagct caaggaggtc ttggtggtaa attacttaca aatttcttac cattgaaagg 420
ccagaaacga ataattcacc ttgatctaaa acttatagct gatgtaggcc tagtaggatt 480
cccaaatgct ggaaaatcct ctttgctaag ttgtgtttct catgcaaaac ctgcaattgc 540
agattacgca tttacaacat taaagctgaa gctcggaaag ataatgtaca gtgatttcaa 600
acagatatca gtagctgatc ttccgggttt aatagaagga gcacatatga acaaaggaat 660
gggccacaaa ttcctcaagc atatagaaag aactagacaa ctacttttg ttgttgatat 720
ttctggattt cagctttctt ctcacactca atacaggaca gcttttgaaa ccataatact 780
gcttacaaaa gagttggaat tgtacaaaga ggaacttcag acaaaacctg cactcttggc 840
agttaataaa atggacttgc cagatgccca agataagttc catgaattga tgagccagct 900
ccagaatcct aaagattttc tgcatttatt tgaaaaaaac atgattccag agaggactgt 960
agagttccaa catatcatcc ccatatctgc agttactgga gaaggaatcg aagaattaaa 1020
gaattgtata agaaagtcac tggatgaaca ggccaaccag gaaaatgatg cacttcataa 1080
gaaacagttg cttaatttgt ggatttctga tacaatgtct tctactgagc caccatcaaa 1140
gcatgctgtt actacttcca aaatggatat aatttaaata tattaaaaat ggtattgatg 1200
gaacagtaaa aaaaaaaa 1218

<210> 116
<211> 1286
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2111754CB1

<400> 116
cccgccttga acttctggac tagcccctcg attgttgtag atgccaagcg gacctcgcgc 60
cgctctgcgt tgggccagcc cctcacagct ggtttcttac cacgtattgc gcaacggaat 120
ctatgcctgt tacccacact ccctgcgccc ccgcaccccg ctcctgtgcg caagtcggaa 180
tataaaaccg cggaggagtg agctcttggg gtgtccagtt ggttgccgcg gcagtctctc 240
cgagcagcgc atttgtcttc taggctgctt ggttcgtgcc tccgagaaag gggtctcctg 300
ctgccagcta agtgtgggag aacttgtgca cgtatctccc ctccgaatcc caacgatggg 360
taacgccagc tttggctcca aggaacagaa gctgctgaag cggttgcggc ttctgcccgc 420
cctgcttatc ctccgcgcct tcaagcccca caggaagatc agagattacc gcgtcgtggt 480
agtcggcacc gctggtgtgg ggaaaagtac gctgctgcac aagtgggcga gcggcaactt 540
ccgtcatgag tacctgccga ccattgaaaa tacctactgc cagttgctgg gctgcagcca 600

cggtgtgctt tccctgcaca tcaccgacag caagagtggc gacggcaacc gcgctctgca 660
gcgccacgtt atagcccggg gccacgcctt cgtcctggtc tactcagtca ccaagaagga 720
aaccctggaa gagctgaagg ccttctatga gctgatctgc aagatcaaag gtaacaacct 780
gcataagttc cccatcgtgc tggtgggcaa taaaagtgat gacacccacc gggaggtggc 840
cctgaatgat ggtgccacct gtgcgatgga gtggaattgc gccttcatgg agatttcagc 900
caagaccgat gtgaatgtgc aggagctgtt ccacatgctg ctgaattaca agaaaaagcc 960
caccaccggc ctccaggagc ccgagaagaa atcccagatg cccaacacca ctgagaagct 1020
gcttgacaag tgcataatca tgtgagccct gggccttaag agccagctct tcctatcctg 1080
tagcgtgtag aaaacgtgga ctcatttcac tatgttacat gtacatggtt gattttgtgc 1140
tgttgtttgg actgtaacat ccatgttgtc aatacgtata ccttgtaagt ggataacttt 1200
tctttttccc aggccagaga attcaaattg ttaaaacatt ggcatttgaa gaggagaaca 1260
aaatgtagca tgatgtattt aaagta 1286

<210> 117
<211> 3057
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2123286CB1

<400> 117
caaggctccg gcctgcgagg agtcacatta actttgctct agaagacaac tttacaagga 60
tctaaaagga acaggattaa agatgactga atactgggtt ccagaaattt aaaacaatca 120
gcttagcaaa tcatatattc ttctgtggag ctgagaattg atgtccgctc ttccccgtga 180
tttggaactt tccaatccca gagaaaagtt gacaaaggga ctgcccagga ctgagtccat 240
atggaagaag aacttcctct tttctctgga gacagtggca agccagtaca ggctactctg 300
tcatctttga agatgttaga tgtgggaaag tggccaattt tttccctttg ttctgaagaa 360
gaactacagt taattcgtca ggcttgtgtc tttggcagtg ctggcaatga agttttatac 420
actacagtaa atgatgagat ttttgtgctt ggcacaaact gctgtggctg tttggggtta 480
ggtgacgtcc agagcaccat tgaacctcgg agactggatt ctttaaatgg caaaaaaata 540
gcctgcctca gctatgggag tggtccacat attgtccttg caacaacaga aggagaagtc 600
tttacctggg gtcataatgc ttatagccag ctgggcaatg ggacaactaa tcatggttta 660
gtgccctgtc atatctctac taatctgtca aacaaacaag tcattgaagt tgcctgtggg 720
tcttaccatt ctttggtgct aacatctgat ggagaggtat ttgcctgggg ttataataac 780
tctgggcagg taggatctgg atcaacagtt aatcagccaa tccctcgaag agtcactggc 840
tgcctacaaa ataaagtagt tgtgaccata gcatgtgggc agatgtgctg catggcagta 900
gtagacacgg gggaggtcta tgtctggggt tacaacggaa acgggcagct tggactcggc 960
aacagtggca accagccaac cccttgcaga gtggcagctt tgcaaggcat ccgtgtccag 1020
agggtcgcct gtggctacgc acacacatta gtattaacag atgaaggcca agtgtatgct 1080
tggggcgcca attcttatgg gcagttgggc actggcaata aaagcaacca gtcctatcct 1140
actcctgtca ctgtggaaaa ggacaggatt atcgagattg cagcctgtca ctccacacac 1200
acgtctgcgg ccaagacgca gggtgggcac gtgtacatgt ggggccagtg ccggggtcag 1260
tccgtgatcc tcccgcacct cacccacttc tcctgcactg acgacgtgtt tgcctgcttt 1320
gccacgcccg ccgtcacgtg gcgcctcctc tccgtggaac ctgatgacca cctcacagtg 1380
gctgagtcac tgaagaggga atttgacaac ccggacactg cagacctgaa gtttctagtt 1440
gatggaaagt acatttatgc acataaagtc cttctcaaga ttcgatgtga gcattttcgt 1500
tcgtcattgg aagataacga ggatgatatt gtagaaatga gtgaattttc atatcctgtt 1560
taccgggcct tcctggaata cctatacaca gacagcatca gcctttctcc tgaggaggca 1620
gtaggactgc tagacttggc tacattttat agagaaaatc gtttgaaaaa gctctgccaa 1680
caaactatca agcaaggcat ctgcgaggag aatgccatcg ctctgctctc ggctgcggtg 1740
aagtatgatg cacaggattt agaagaattc tgcttcaggt tttgcataaa ccatctgact 1800
gtagtaacac aaacatcagg ttttgcagaa atggaccatg atctcctgaa gaactttatc 1860
agcaaagcaa gcagagttgg agcctttaaa aattgatccc atctgcagga aagtttttga 1920
gcctttccat ttcccctgca aaagccagag atgaatcact tctctttaat taatagtatg 1980
tatgatgagc tatgtttggc tgagtacttg taactgtcag aagaaggatg gtggtgagtg 2040
gtctttgtct gcctaaaccc agagtttatg tagaaagcat tgaatgttct gatcagatgt 2100
gactaaggtc aaggaaaaaa aattgaaata tcttatttac catttcctct ttttgagtca 2160
cttaaattgg acacctttgg taccctggtc tcagtatatg ctattctggc ccaaatgttc 2220
cattattcag ctggctgata ccacatagat agcttgacaa ggagtgctgt ctgtccttac 2280
cacattttca gcactcagca cagtgccttg tgtataatag gcactcaatt tattataaat 2340
cttcagtatg tcctgagaac agctttagtc atggaatact gggagaagga ataactttca 2400

caaaataaac ttaaaacagc ctgtaattat tgaggttcat attcttctgg tatatcattc 2460
tgagaaattg tggctaattt agaacattgt ttagaattga caaaaggccc tggcaattaa 2520
attgtcaagg cccaagggct aattttaatt ttctttttac ttggagtcat tcattaattt 2580
ctcacatggg attatggagt atgaagtatt atctttgaat gaaattcctg ggctgatctg 2640
ccttacataa tcacataagg tcctttgctt ttctttgtgt taagagggac ttgcctctgt 2700
aaatgaaaat gacaatgtgc ttttcttgta gttgactttc atgtcactca ctataaaata 2760
ggtctcttaa cctggcacca gtataactat aaagcactag ctgagaagga actgatactt 2820
acatttcatg gacagcatta acaagaatga gataaatttg tactttttag atcaaaacaa 2880
attacctaat tgcaaaagag aaactgaaat ggaacatagt ctcagattct tctaatgtgt 2940
atctcacaat gtcatgtaat gtaaaggaaa cccttttgga attagaattc ttgttctgat 3000
gctgaactat ttggtaataa agtgcttatt tgcagataac agaaaaaaaa aaaaaaa 3057

<210> 118
<211> 1803
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2477507CB1

<400> 118
ggtcgcggcc ttcacgggtt cctcggccgt cctgagtccc aaatcccaag ctggagccgt 60
tcagctcccc tccaccgctt agagatttgg ggggctctgg ccccgtcctt gcggaccatt 120
ccgaggggag tccagaggtg aggccgagga acctccctga ctttgcgggg cgcgccgctc 180
ctgcgtctcc tgccagtctc ccttccttct tttcggtcaa caattgaaaa caaaacgagg 240
aacagcagag gagctactgt ataccgagcc ctcagcattg ttcgtaatct ccgcctgcta 300
acagccttgt gaagaaggtg ctattcttct caacacttta cagatgagga cacttgaggt 360
tcggagacgt ggagcctctt gcacagctgc ttaagtggtg gtagagccga gatttgaacc 420
ctcctaacca ttcctttctg ccgcctactg cagctcccag cagagatgat tgaactgttg 480
ctcggggtag ggcccccagg tgtcagtaat taacactgtg gatacctccc atgaggacat 540
gattcacgac gcccagatgg actactatgg cacccgcctg gcaacctgct catcagacag 600
gtccgtcaaa atctttgatg tgcgcaatgg agggcagatc cttatcgccg acctcagggg 660
tcatgagggt cctgtgtggc aagtggcctg ggctcacccc atgtacggca acatcctggc 720
atcgtgctcc tatgaccgga aagtcattat ctggagagag gaaaacggca cctgggagaa 780
gagccacgag catgcgggac acgactcctc agtgaactcg gtgtgctggg ccccccatga 840
ctacggcctg atcctggcct gtgggagctc ggatggggcc atctccctgc tgacttacac 900
cggggaaggc caatgggaag taaagaagat caacaacgct cacaccattg gctgcaatgc 960
cgtcagctgg gcccctgctg ttgtacctgg aagcctcata gaccacccat cggggcagaa 1020
acccaattac atcaagaggt ttgcatcagg tggctgtgac aacctcatca agctgtggaa 1080
ggaggaggag gacggccagt ggaaggagga gcagaagcta gaagcgcaca gtgactgggt 1140
tcgagatgtg gcctgggccc cctccatcgg cctgccccacc agcaccatcg ccagctgctc 1200
ccaggatggt cgtgtgttca tttggacctg tgatgatgcc tcaagcaata cgtggtcccc 1260
taaattgttg cacaagttca acgatgtggt gtggcatgtg agctggtcca tcacagccaa 1320
catcctggct gtctctggtg gagacaataa ggtgaccctg tggaaggagt cagttgatgg 1380
gcagtgggtg tgcatcagtg atgtcaacaa gggccagggc tccgtatcag catcagtgac 1440
agagggccag cagaacgagc agtgacaaga caggtggggc ctggctcccc acccgccagc 1500
tccaggactg ccccttcctg ggccaactaa ccagacaact gggaagagcc cccaactcca 1560
acaggattat tttcccagga ggagttacag atgcagccac agattgatca tctgccttaa 1620
cgtgatcgga gatgctttgt aatctactgt ccagctgaaa gcactcatgt tacgaggaag 1680
aaactacaag tgatgttcaa atctattttg ggtcattttt atgtaccttt gggttcaggc 1740
attatttggg gggttttgtt tccaaaggaa ctaaataaag tcatattgct tataaaaaaa 1800
aaa 1803

<210> 119
<211> 4407
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2759119CB1

<220>
<221> unsure
<222> 4373, 4379
<223> a, t, c, g, or other

<400> 119
ggtcgtcatt ggacaaccgc cgcggggccc tggtctctgc tacctgtagc tgagggtgct 60
gttgatgggc agcgcggcgc gctgggaagg ctcgttctcg cgagagttca gctcccttct 120
atacccgtgg ctgcctcagc acctcgagga tcgacatgga cgctctcgag gactacgttt 180
ggccgcgggc aacctcggag cttatactcc tcccagtgac gggtctggag tgcgtggggg 240
accggctgtt ggcgggtgag ggtcccgatg tcctggtgta cagcttggac tttggtgggc 300
atctgcggat gataaagcga gtgcagaacc tgcttggcca ctatcttatc catggcttcc 360
gggtacggcc agagcctaat ggagaccttg acttggaggc catggtggct gtgtttggaa 420
gcaagggact ccgagttgtg aaaattagct ggggacaggg ccacttctgg gagctttggc 480
gctctggcct gtggaacatg tctgactgga tttgggatgc acgctggctt gagggaaata 540
tagccttggc cctgggccac aactcagtgg tgctatatga ccctgtagta gggtgcatcc 600
tgcaagaggt gccctgcaca gacaggtgca ccctctcttc agcctgcctg attggagacg 660
cctggaagga gctgaccata gtggcaggtg ctgtttccaa ccagctcttg gtctggtacc 720
cagcaactgc cttagcagac aacaaacctg tagcacctga ccgacgaatc agtgggcatg 780
tgggcatcat cttcagcatg tcatacctgg aaagcaaggg attgctggct acagcttcag 840
aagaccgaag cgttcgtatc tggaaggtgg gcgacctgcg agtgcctggg ggtcgggtgc 900
agaatattgg gcactgcttt gggcacagcg cccgtgtgtg gcaggtcaag cttctagaga 960
attaccttat cagtgcagga gaggattgtg tctgcttggt gtggagccat gaaggtgaga 1020
tcctccaggc ctttcgggga caccagggac gtgggatccg ggccatagct gcccatgaga 1080
ggcaggcctg ggtgatcact ggggggtgatg actcaggcat tcggctgtgg cacttggtag 1140
ggcgtgggta ccggggattg gggtctcgg ctctctgctt caagtcccgt agtaggccag 1200
gtacactcaa ggctgtgact ctggctggct cttggcgact gctggcagtg actgatacag 1260
gggccctgta tctctatgac gtcgaggtca agtgctggga gcagctgcta gaggataaac 1320
atttccagtc ctactgcctg ctggaggcag ctcctggtcc cgagggcttc ggattgtgtg 1380
ctatggccaa tggggaaggt cgtgtcaagg ttgtccccat caacactcca actgctgctg 1440
tggaccagac cctgtttcct gggaaggtgc acagcttgag ctgggccctg cgtggttatg 1500
aggagctcct gttgctggca tcgggccctg gcggggtagt agcttgccta gagatctcag 1560
ccgcaccctc tggcaaggcc atctttgtca aggaacgttg tcggtacctg ctgcccccaa 1620
gcaagcagag atggcacaca tgcagtgcct tcctaccccc aggtgacttc ctggtgtgtg 1680
gtgaccgccg gggctctgtg ctgctattcc cctccagacc aggtctgctc aaggaccctg 1740
gggtgggagg caaggctcgg gctggtgctg gggcacctgt agtgggtagt ggtagtagtg 1800
ggggtgggaa tgctttcact gggttgggcc cagtgtctac cctgccctct ctgcacggga 1860
agcagggtgt gacctcagtc acatgccatg gtggctatgt gtataccata gggcgtgatg 1920
gagcctacta ccagctgttt gtacgagacg gccagctcca gccagtccta aggcagaagt 1980
cctgtcgagg catgaactgg ctagctgggc tccgtatagt gccccgatggg agcatggtta 2040
tcctgggttt ccatgccaat gagtttgtgg tgtggaaccc tcggtcacac gagaagctgc 2100
acatcgtcaa ctgtggtgga gggcaccgtt cgtgggcatt ctctgatact gaggcggcca 2160
tggcctttgc ttacctcaag gatggggatg tcatgctgta cagggctctg ggtggctgca 2220
cccggccaca cgtgattctc cgggagggtc tgcatggccg tgagatcact tgtgtaaagc 2280
gtgtgggcac cattaccctg ggcctgaat atggagtgcc cagcttcatg cagcctgatg 2340
acctggagcc tggcagtgag gggcccgact tgactgacat tgtgatcaca tgtagtgagg 2400
acactactgt ctgtgtccta gcactcccta caaccacagg ctcagcccac gcactcacag 2460
ctgtttgtaa ccatatctcc tcggtacgtg ctgtggctgt gtggggcatt ggcaccccag 2520
gtggccctca ggatcctcag ccaggcctga ctgcccatgt ggtgtctgcg ggggggcggg 2580
ctgagatgca ctgcttcagc atcatggtta ctccggaccc cagcacccca agccgcctcg 2640
cctgccatgt catgcacctt tcgtcccacc ggctagatga gtattgggac cggcaacgca 2700
atcggcatcg gatggttaag gtagacccag agaccaggta catgtccctt gctgtgtgtg 2760
aacttgacca gcccggcctt ggccccccttg tggctgcagc ctgtagtgat ggggccgtaa 2820
gctctttctt ttgcaggatt ctgggcggat tctgcagctc cttgctgaaa ccttccacca 2880
taagcgatgt gtcctcaagg tccactcctt tacacacgag gcacccaacc agaggcggag 2940
gctcctcctg tgcagcgcag ctactgatgg cagcctggct ttctgggatc tcaccaccat 3000
gctagaccat gactccactg tcctggagcc tccagtggat cctgggcttc cctaccggct 3060
tggcaccccc tccctgactc tccaggccca cagctgtggt atcaacagcc tgcacacctt 3120
gcccacccgt gagggccacc atctcgtggc cagtggcagt gaagatggat ccctccatgt 3180
cttcgtgctt gctgtggaga tgctacagct agaagaggct gtgggagagg ctgggctggt 3240
accccagctg cgtgtgctag aggaatactc tgtcccctgt gcacatgctg cccatgtgac 3300
aggcctcaag atcctaagcc caagcatcat ggtctcagcc tccattgatc aacggctgac 3360
cttctggcgt ctggggcatg gtgaacccac cttcatgaat agcactgtgt tccatgtgcc 3420

tgatgtggct gacatggact gctggcctgt gagccctgag tttggccacc gttgtgccct 3480
tgggggtcag gggcttgagg tttacaactg gtatgactga ggtatcctgc ggtggctggc 3540
gtgctgggca tggggcctgc tcacagacag catggagcag ggaagggctg tctgtgccca 3600
tgctcagcat gccttgaggg gaggaggtgg tggccgtggg ttcctgatgt cggtgcagga 3660
gctgaaggtg agtggagtgc tgccaagaat atgcccgact ccccatgaca agacagaact 3720
ttgtaacaaa cagtaccaat ttattttggc cgtgggtttt tgcttttttt ccagttgatg 3780
actttgtgaa cattcccagg tattggagcc tctgtggcct taaatgtggc tcagtggagg 3840
gagacccagc atagccaggc cagtatggag cacctcacgc acagctctca gaagctgcag 3900
gcggacgaac atctgaccaa agaggtgtgg tcgaggctcc tgaaagagaa agggcctgct 3960
ggtctcatcc tctgcttcct ttgcctttac cctatacctc tctgcacgtc ccaccccgtt 4020
ttgctgtgtg ctcaccccca ggatgtgtac ccggttgtag taggagctga aatccatgct 4080
gagctgtacc aggaacttgc atatctagag acagagactg agtcactggc ccatctcttt 4140
gctcttgtgc cccaggccag aataaagaat agagtgtaga gtgtcctggt tgtctatgcc 4200
tcaccatctc tgtgcgtaca gcaatgtgga ccccggggct gtgcagtcca gcactgctgt 4260
ccggctcagc agatccggaa agggaggata ctgttgaaga gcaacaacca ctcaccctgt 4320
ttggggagaa aagtgcttgg aaggggaatc caggctcctt gtgccagtaa cangagggnc 4380
aatcactcat catgtagcag tgagaag 4407

<210> 120
<211> 959
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2823818CB1

<400> 120
cccacgcgtc cgcccacgcg tccgggagcg tggagcgccg ggactgtgca cgcttgaccg 60
gaagcccaga ccagtgcggt cctagccaga gagaaaggac atttgccaac aatgagacac 120
gaagcgccca tgcagatggc ctctgcccaa gatgccaggt acggccagaa agactcctct 180
gatcagaact ttgactacat gttcaaatta ctcatcatcg gcaatagcag tgtggggaaa 240
acatcttttc tattccgtta tgcagatgac tcctttacat ctgcattcgt cagcacagtt 300
gggatcgatt tcaaagtaaa aactgtattc aaaaatgtaa agagaatcaa gcttcagatt 360
tgggacacag caggccagga aagatacagg actatcacca cagcctatta tcgtggagcc 420
atgggcttta ttttaatgta tgacattaca aatgaagaat ccttcaatgc agtacaagat 480
tggtcaactc aaatcaaaac atactcttgg gacaatgccc aagttattct ggttgggaac 540
aagtgtgaca tggaagacga gcgggtcatc tcaactgagc gaggtcaaca tttaggagaa 600
cagcttgggt ttgagttttt tgaaacaagt gccaaggaca acattaatgt caagcagaca 660
tttgagcgcc ttgtggatat catctgcgac aaaatgtcag agagtttgga gactgatcct 720
gccatcactg ctgcaaagca gaacacgaga ctcaaggaaa ctcctcctcc accgcagccc 780
aactgtgcct gctagtgtcc ccgtgcacac aggcagctcc agggggctct ggttgccaac 840
aaacagcatt tgtaaatggt ctattagcct tcatttatac tgcctaacaa ttatttgaag 900
gaataaattg atgtcaatgg ctcgtaaaaa aaaaaaaaaa aagtaaaaaa aaaaaaaaa 959

<210> 121
<211> 1809
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2859730CB1

<400> 121
ggcagcggtt ggaggcttcg cccggctttg cagcggggac ttcggcggcg gcgcctcagg 60
cacctcggcc cggacacgat gaggcgagtg gtacgacaga gcaagtttcg gcatgtattt 120
gggcaagcgg tgaaaaatga ccagtgctat gatgacatcc gggtttctcg tgtgacctgg 180
gatagttcct tttgtgctgt caatcccaga tttgttgcca taatcataga ggcaagtggg 240
ggaggagcgt tccttgtcct ccctctgcgc aagactggtc gaattgacaa atcttaccct 300
acagtatgtg gccacacagg accagtgctg gacatagact ggtgcccaca taacgatcag 360
gtcattgcca gcggttcaga ggactgcacg gtcatggtat ggcagatccc agaaaatgga 420
ctcacccttt ccctgactga acctgtggtg attttggaag gccactcaaa gagagtcggc 480

atcgtggctt ggcatccaac ggcccgcaat gtgcttctta gtgcaggctg tgataatgcc 540
attatcatct ggaatgtggg aacaggggaa gcccttataa acttggacga tatgcattca 600
gacatgattt acaatgtgag ctggaaccgg aatggcagtc tgatctgcac agcttccaaa 660
gacaagaaag tgagagtcat tgatcccagg aaacaagaga ttgttgctga gaaggagaaa 720
gcacatgaag gagcaagacc catgagagcc atcttcctgg ccgatggcaa tgtcttcacc 780
actgggttca gccgcatgag cgagcggcag ctggctctct ggaatccgaa aaatatgcag 840
gaaccaattg ctcttcatga gatggacact agcaatgggg tgttgctgcc tttctatgac 900
cctgacacca gcatcattta cttatgtgga aagggtgaca gcagtattcg ctattttgag 960
atcacggatg aatccccgta cgtccactac ctcaacacat tcagcagcaa ggagcctcag 1020
agagggatgg gttacatgcc caagagggga cttgatgtta acaaatgtga gattgccaga 1080
ttcttcaaac ttcatgagag aaagtgtgaa cctattatta tgactgttcc caggaagtct 1140
gaccttttcc aagatgacct gtatcctgac acagcggggc cagaggccgc gctggaggca 1200
gaagagtggt tcgaaggcaa gaatgcagac ccaatcctca tctccttgaa gcacgggtac 1260
attccaggca aaaacaggga tctcaaggtg gtcaagaaga acattctgga tagcaagccc 1320
actgcaaaca agaagtgcga cctgatcagc atccccaaga aaaccacaga cacggccagt 1380
gtgcaaaatg aagccaagtt ggatgagatt ttaaaagaga tcaaatctat aaaagacaca 1440
atctgcaatc aagatgagcg tatttccaag ttagaacagc agatggcaaa gatagcagcc 1500
tgaaggtccc accccccacc ctacagaaaa aatgggagca agaacttgtg cttgggagct 1560
ggttattggt gtggtcctag ggagggcgga aagggaggca ctgccatttg gagacattcc 1620
atttcagatt tgtcaaccag cgataggcca cattccagta agaactcaat ttgtctccca 1680
aatttgcaga aacaaaacgt gatttaaaag ctgagctttt tatcagaaag ctttttttgat 1740
gttttaagtg ttatgtgact tgttgaactt tttaaaaagt gctactttta aaatcccaga 1800
tactctgaa 1809

<210> 122
<211> 2028
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 2861155CB1

<220>
<221> unsure
<222> 1943, 2003
<223> a, t, c, g, or other

<400> 122
tggcgggttc cgtgggtcgc ccgcgaaatc tgatccggga tgcggcggcc caatcggaag 60
gtggaccgaa atcccgcgac agcaagaggc ccgtagcgac ccgcggtgct aaggaacaca 120
gtgctttcaa aagaattggc gtccgctgtt cgcctctcct cccgggagtc ttctgcctac 180
tcccagaaga ggagggaagc acaggtgggt ttctttagct ctgcgtcgga tccctgagaa 240
cttcgaagcc atcctggctg aggctaatct ccgctgtgct tcctctgcag tatgaagact 300
ttggagactc aaccgttagc tccggactgc tgtccttcag accaggaccc agctccagcc 360
catccttctc cccacgcttc cccgatgaat aaaaatgcgg actctgaact gatgccaccg 420
cctcccgaaa gggggggatcc gccccggttg tccccagatc ctgtggctgg ctcagctgtg 480
tcccaggagc tacgggaggg ggacccagtt tctctctcca ctcccctgga aacagagttt 540
ggttccccta gtgagttgag tcctcgaatc gaggagcaag aactttctga aaatacaagc 600
cttcctgcag aagaagcaaa cgggagcctt tctgaagaag aagcgaacgg gccagagttg 660
gggtctggaa aagccatgga agatacctct ggggaacccg ctgcagagga cgagggagac 720
accgcttgga actacagctt ctcccagctg cctcgatttc tcagtggttc ctggtcagag 780
ttcagcaccc aacctgagaa cttcttgaaa ggctgtaagt gggctcctga cggttcctgc 840
atcttgacca atagtgctga taacatcttg cgaatttata acctgccccc agagctgtac 900
catgaggggg agcaggtgga atatgcagaa atggtccctg tccttcgaat ggtggaaggt 960
gataccatct atgattactg ctggtattct ctgatgtcct cagcccagcc agacacctcc 1020
tacgtggcca gcagcagccg ggagaacccg attcatatct gggacgcatt cactggagag 1080
ctccgggctt cctttcgcgc ctacaaccac ctggatgagc tgacggcagc ccattcgctc 1140
tgcttctccc cggatggctc ccagctcttc tgtggcttca accggactgt gcgtgttttt 1200
tccacggccc ggcctggccg agactgcgag gtccgagcca catttgcaaa aaagcaggc 1260
cagagcggca tcatctcctg catagccttc agcccagccc agcccctcta tgcctgtggc 1320
tcctacggcc gctccctggg tctgtatgcc tgggatgatg gctcccctct cgccttgctg 1380
ggagggcacc aagggggcat cacccacctc tgctttcatc ccgatggcaa ccgcttcttc 1440

tcaggagccc gcaaggatgc tgagctcctg tgctgggatc tccggcagtc tggttaccca 1500
ctgtggtccc tgggtcgaga ggtgaccacc aatcagcgca tctacttcga tctggacccg 1560
accgggcagt tcctagtgag tggcagcacg agcggggctg tctctgtgtg ggacacggac 1620
gggcctggca atgatgggaa gccggagccc gtgttgagtt ttctgcccca gaaggactgc 1680
accaatggcg tgagcctgca ccctagcctg cctctcctgg ccactgcctc cggtcagcgt 1740
gtgtttcctg agcccacaga gagtggggac gaaggagagg agctgggcct tcccttgctc 1800
tccacgcgcc acgtccacct tgaatgtcgg cttcagctct ggtggtgtgg gggggggcca 1860
gactccagca tccctgatga tcaccagggc gagaaagggc agggaggaac aggagggagg 1920
tcgtgggggg cgtgatataa aanggtgttt gagtggctgt gactccttcc tacacagggc 1980
cctgataagc ctaggaatgc canagcccag ctgtagggtc ccagtccc 2028

<210> 123
<211> 2223
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3002667CB1

<400> 123
gcgcgcacgt ggggccgggg cggagagagg cgagcaccgg gaaggggagc gtggggccgc 60
tggaatgggt gaatttaagg tccatcgagt acgtttcttt aattatgttc catcaggaat 120
ccgctgtgtg gcttacaata accagtcaaa cagattggct gtttcacgaa cagatggcac 180
tgtggaaatt tataacttgt cagcaaacta ctttcaggag aaatttttcc caggtcatga 240
gtctcgggct acagaagctt tgtgctgggc agaaggacag cgactcttta gtgctgggct 300
caatggcgag attatggagt atgatttaca ggcgttaaac atcaagtatg ctatggatgc 360
ctttggagga cctatttgga gcatggctgc cagccccagt ggctctcaac ttttggttgg 420
ttgtgaagat ggatctgtga aactatttca aattacccca gacaaaatcc agtttgaaag 480
aaattttgat cggcagaaaa gtcgcatcct gagtctcagc tggcatccct ctggtaccca 540
cattgcagct ggttccatag actacattag tgtgtttgat gtcaaatcag gcagcgctgt 600
tcataagatg attgtggaca ggcagtatat gggcgtgtct aagcggaagt gcatcgtgtg 660
gggtgtcgcc ttcttgtccg atggcactat cataagtgtg gactctgctg ggaaggtgca 720
gttctgggac tcagccactg ggacgcttgt gaagagccat ctcatcgcta atgctgacgt 780
gcagtccatt gctgtagctg accaagaaga cagtttcgtg gtgggcacag ccgagggaac 840
agtcttccat tttcagctgg tccctgtgac atctaacagc agtgagaagc agtgggtgcg 900
gacaaaaccg ttccagcatc acactcatga cgtgcgcact gtggcccaca gcccaacagc 960
gctgatatct ggaggcactg acacccactt agtctttcgt cctctcatgg agaaggtgga 1020
agtaaagaat tacgatgccg ctctccgaaa aatcaccttt ccccaccgat gtctcatctc 1080
ctgttctaaa aagaggcagc ttctcctctt ccagtttgct catcacttag aactttggcg 1140
actgggatcc acagttgcaa caggcaagaa tggggatact cttccactct ctaaaaatgc 1200
agatcattta ctgcacctaa agacaaaggg tcctgagaac attatctgta gctgtatctc 1260
cccatgtgga agttggatag cctattctac agtttctcgg ttttttctct atcggctgaa 1320
ttatgaacat gacaacataa gcctcaaaag ggtttccaaa atgccagcat tccttcgctc 1380
tgcccttcag attttgtttt ctgaagattc aacaaagctc tttgtagcat caaatcaagg 1440
agctctgcat attgttcagc tgtcaggagg aagcttcaag cacctgcatg ctttccagcc 1500
tcagtcagga acagtggagg ccatgtgtct tttggcagtc agtccagatg ggaattggct 1560
agctgcatca ggtaccagtg ctggagtcca tgtctacaac gtaaaacagc taaagcttca 1620
ctgcacggtg cctgcttaca atttcccagt gactgctatg gctattgccc ccaataccaa 1680
caaccttgtc atcgctcatt cggaccagca ggtatttgag tacagcatcc cagacaaaca 1740
gtatacagat tggagccgga ctgtccagaa gcagggcttt caccaccttt ggctccaaag 1800
ggatactcct atcacacaca tcagttttca tcccaagaga ccgatgcaca tccttctcca 1860
tgatgcctac atgttctgca tcattgacaa gtcattgccc cttccaaatg acaaaacctt 1920
actctacaat ccatttcctc ccacgaatga atcagatgtc atccggaggc gcacagctca 1980
tgcttttaaa atttctaaga tatataagcc tctactcttc atggatcttt tggatgaaag 2040
aacactcgtg gcagtagaac ggcctctgga tgacatcatt gctcagctcc caccacccat 2100
taaaaagaag aaatttggaa cctaaaacag ggcactgtct gtgtccttcc ttgaactgtc 2160
taccctgttg cttttcacaa atcatggtaa taaaacaagt tattcttgag gaaaaaaaaa 2220
aaa 2223

<210> 124
<211> 728
<212> DNA

<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3043734CB1

<400> 124
gcggcgtttc tggtggccag gcatcccggt cctcgcgcgt ggcgcagctc ccatcgccgg 60
accgacccat gtcgcgcccg cattgggtcc cgggaccccg gcgggagtgc cgcgtccgtc 120
ctttccagtc gccgggagtc tgagtcgcgg gccacgcggg agtggcggtg gagagcccgc 180
cggtcgttat gaggacggat ctaaaatgac cagcaaacgg aaaccttgcc aaacgcagct 240
caggagatcc atcagtgagc agttgcggga ctccacggcc agagcctggg atctgctgtg 300
gaagaacgtc cgggagaggc ggctggcaga aattgaggca aaagaagcat gtgactggct 360
ccgtgctgcc gggttcccgc aatacgctca gttatatgag gattcacaat ttcccatcaa 420
cattgtggct gtcaagaatg atcatgattt tcttgaaaag gaccttgtag aacctctttg 480
caggtaaacc atgtgaagta tttttgtttc tttccactgt tcagtctgca acaggcatca 540
ctatactgaa gggcgagctc agctattcgg caagtattca ctgagtgcct accatgtgcc 600
tgacccaggt gcaggttcta aatgtactac tgttaatgag catgatcagt ttgtgttttc 660
atggagctta aatcctagca ggggcctttg gacactagat taggaaaatg acagagaaag 720
aagagaga 728

<210> 125
<211> 2161
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3294893CB1

<400> 125
gaggcggaag agcttctcgg ctctaggctc tggagtcccg ggagcagtga ggggccaccc 60
ggggcacagg aaagggccgc taggggaggg ccgggtgcac tcggggtgtc tggccgcgg 120
gtctgaggga tgaggagggg ccatggccag cgacggggcc aggaagcaat tctggaagcg 180
cacaacagca agctcccggg cagcatccag cacgtgtatg gtgcccagca ccccccccttt 240
gatccactgt tacatggcac tttgctcagg tccacggcca agatgccgac cacaccagtg 300
aaggccaaga gggtcagcac cttccaggag tttgagagca ataccagcga tgcctgggac 360
gctggggagg acgacgatga gctcctggcc atggcggcgg agagcctgaa ctccgaggtg 420
gtcatggaga cggccaaccg tgtgctgcgt aaccacagcc agcggcaggg gcggcccacg 480
ctgcaggagg ggccagggct tcagcagaag cccaggcccg aggcagagcc gccctcaccc 540
cccagcggcg acctccggct ggtgaagtcg gtcagtgaga gccacacgtc ctgtcctgca 600
gaaagtgcca gcgatgccgc ccctctgcag aggtcccagt ctctcccaca ctcggccacc 660
gtcacgctgg gtggcacatc tgaccccagc actctcagca gctcagcgct gagcgaaaga 720
gaggcctccc ggctcgacaa gttcaagcag ctgcttgccg gccccaacac ggaccttgag 780
gaattacgga ggttgagctg gtccggaatc cctaagccag tgcgtccaat gacgtggaag 840
ctcctctcag gttaccttcc cgccaatgta gaccggagac cagccactct ccagagaaaa 900
caaaaagaat attttgcatt tattgagcac tattacgatt ctaggaacga cgaagttcac 960
caggacacat acaggcagat ccacatagac atccctcgca tgagccctga agcgttgatc 1020
ctgcagccca aggtgacgga gatttttgaa aggatcttgt tcatatgggc gatccgccac 1080
ccagccagtg gatacgttca gggtataaat gatctcgtca ctcctttctt tgtggtcttc 1140
atttgtgaat acatagaggc agaggaggtg gacacggtgg acgtctccgg cgtgcccgca 1200
gaggtgctgt gcaacatcga ggccgacacc tactggtgca tgagcaagct gctggatggc 1260
attcaggaca actacacctt tgcccaacct gggattcaaa tgaaagtgaa aatgttagaa 1320
gaactcgtga gccggattga tgagcaagtg caccggcacc tggaccaaca cgaagtgaga 1380
tacctgcagt ttgccttccg ctggatgaac aacctgctga tgagggaggt gcccctgcgt 1440
tgtaccatcc gcctgtggga cacctaccag tctgaaccgg acggcttttc tcatttccac 1500
ttgtacgtgt gcgctgcttt tctcgtgaga tggaggaagg aaatactaga agaaaaagat 1560
tttcaagagc tgctgctctt cctccagaac ctgcccacag cccactggga tgatgaggac 1620
atcagcctgt tgctggccga ggcctaccgc ctcaagtttg cttttgccga cgcccccaat 1680
cactacaaga aatgagccca ggcccacccg cagctggcct cactgtcccg ggtggcgcgc 1740
cccacctgcc tggctggtgg taggcccctg tgagctggtc ccgggctgct aaaaggcctt 1800
gtgaggtggc ccccaccctcc aggggagctg gtgaagatgg gccacagacc tggtctaggg 1860
ctgacaaaga cagggacagc ctttgttttc tgagatacca aagagagcca ggggagggcc 1920

```
ccgggttcgg cggccagagg caggtcaggg gtcccctctc cctctccctg caatgtcctt 1980
gccaaatgac tgcctcctgc tgcccctagt ccggggcagc ctaggaggcc caccctcttt 2040
ggagtcctgc tgtctgggtg ccagggccgg aacgaggtag tggccatctc atacctactc 2100
tgaaatgcaa aacttctatt ctgttgagtg aaaaaataaa atgtagacaa aaaaaaaaaa 2160
c                                                                 2161
```

<210> 126
<211> 2782
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3349052CB1

<400> 126

```
attagctgcc ggcgtgactt tgaccgcttc ccggtgcgtt accggcagct gaacccaccc 60
ggcgtcacgg gactttgacg cgtgctctgc gcttgccatg agactcctgg gagccgcagc 120
cgtcgcggct ctggggcgcg gaagggcccc cgcctcccta ggctggcaga ggaagcaggt 180
taattggaag gcctgccgat ggtcttcatc aggggtgatt cctaatgaaa aaatacgaaa 240
tattggaatc tcagctcaca ttgattctgg gaaaactaca ttaacagaac gagtccttta 300
ctacactggc agaattgcaa agatgcatga ggtgaaaggt aaagatggag ttggtgctgt 360
catggattcc atggaactag agagacaaag aggaatcact attcagtcag cagccactta 420
caccatgtgg aaagatgtca atattaacat tatagatact cctgggcatg tggacttcac 480
aatagaagtg gaaagggccc tgagagtgtt ggatggtgca gtccttgttc tctgtgctgt 540
tggaggggta cagtgccaga ccatgactgt caatcgtcag atgaagcgct acaacgttcc 600
gtttctaact tttattaaca aattggaccg aatgggctcc aacccagcca gggccctgca 660
gcaaatgagg tctaaactaa atcataatgc agcgtttatg cagataccca tgggtttgga 720
gggtaatttt aaaggtatta tagatcttat tgaggaacga gccatctatt ttgatggaga 780
ctttggtcag attgttcgat atggtgagat tccagctgaa ttaagggcgg cggccactga 840
ccaccggcag gagctaattg aatgtgttgc caattcagat gaacagcttg gtgagatgtt 900
tctggaagaa aaaatcccct cgatttctga tttaaagcta gcaattcgaa gagctactct 960
gaaaagatca tttactcctg tatttttggg aagcgccttg aagaacaaag gagttcagcc 1020
tctttttagat gctgttttag aatacctccc aaatccatct gaagtccaga actatgctat 1080
tctcaataaa gaggatgact caaaagagaa aaccaaaatc ctaatgaact ccagtagaga 1140
caattcccac ccatttgtag gcctggcttt taaactggag gtaggtcgat ttggacaatt 1200
aacttatgtt cgcagttatc agggagagct aaagaagggt gacaccatct ataacacaag 1260
gacaagaaag aaagtacggt tgcaacggct ggctcgcatg catgccgaca tgatggagga 1320
tgttgaggaa gtatatgccg gagacatctg tgcattgttt ggcattgact gtgctagtgg 1380
agacacattc acagacaaag ccaacagcgg cctttctatg gagtcaattc atgttcctga 1440
tcctgtcatt tcaatagcaa tgaagccttc taacaagaac gatctggaaa aattttcaaa 1500
aggtattggc aggtttacaa gagaagatcc cacatttaaa gtatactttg acactgagaa 1560
caaagagaca gttatatctg gaatgggaga attacacctg gaaatctatg ctcagaggct 1620
ggaaagagag tatggctgtc cttgtatcac aggaaagcca aaagttgcct ttcgagagac 1680
cattactgcc cctgtcccgt ttgactttac acataaaaaa caatcaggtg gtgcaggcca 1740
gtatggaaaa gtaataggtg tcctggagcc tctggaccca gaggactaca ctaaattgga 1800
attttcagat gaaacattcg gatcaaatat tccaaagcag tttgtgcctg ctgtagaaaa 1860
ggggttttta gatgcctgcg agaagggccc tctttctggt cacaagctct ctgggctccg 1920
gtttgtcctg caagatggag cacaccacat ggttgattct aatgaaatct ctttcatccg 1980
agcaggagaa ggtgctctta aacaagcctt ggcaaatgca acattatgta ttcttgaacc 2040
tattatggct gtggaagttg tagctccaaa tgaatttcag ggacaagtaa ttgcaggaat 2100
taaccgacgc catggggtaa tcactgggca agatggagtt gaggactatt ttacactgta 2160
tgcagatgtc cctctaaatg atatgtttgg ttattccact gaacttaggt catgcacaga 2220
gggaaaggga gaatacacaa tggagtatag caggtatcag ccatgtttac catccacaca 2280
agaagacgtc attaataagt atttggaagc tacaggtcaa cttcctgtta aaaaaggaaa 2340
agccaagaac taactttgtt tactgtgagt tgactgactc taattgaatc tgcgtggttt 2400
tgatactttg atggattcca gtggaataaa ttcaggctgc tgaaacaaga aattctgagc 2460
ccaggaagcg ggctcttctt tcttcaaaag aagcccttct tgttcatatt caggagcttc 2520
tgttatattc aaaggtaatt ctatgtctat ctcaactcta ttgattggtt ttatagttta 2580
ttgaaaatcc tcaaataaaa tataattatt actgaaatat gtttaatatt taaggggaaa 2640
agagactaat ttcagttata cttttaagct tagaatgtat gttcatttcc aaattttgta 2700
tcataagagt tttcaacata gagaaaagct gaaaaaatgc aaagaataac cacatacttt 2760
ccatctacct tcctttggta ac                                          2782
```

<210> 127
<211> 3019
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3357264CB1

<220>
<221> unsure
<222> 985
<223> a, t, c, g, or other

<400> 127
tggctgggtc cgcgggcggg ggaaggtgtc ctagcggccc gagcctgcgc tccggattct 60
caggcccatc ctgtggtagg ccgtcccagg caggagttgc ctcggaggat ttggcagcca 120
cgacatccca tcctagcccc gcgatgtgcg gggctgtaat cccettgcac aaaccggccg 180
gacgtaaatt gcagaatcaa agagctgctt tgaatcagca gatcctgaaa gccgtgcgga 240
tgaggaccgg agcggaaaac cttctgaaag tggccacaaa ctcaaaggtg cgggagcaag 300
tgcggctgga gctgagcttc gtcaactcag acctgcagat gctcaaggaa gagctggagg 360
ggctgaacat ctcggtgggc gtctatcaga acacagagga ggcatttacg attcccctga 420
ttcctcttgg cctgaaggaa acgaaagacg tcgactttgc agtcgtcctc aaggatttta 480
tcctggaaca ttacagtgaa gatggctatt tatatgaaga tgaaattgca gatcttatgg 540
atctgagaca agcttgtcgg acgcctagcc gggatgaggc cggggtggaa ctgctgatga 600
catacttcat ccagctgggc tttgtcgaga gtcgattctt cccgcccaca cggcagatgg 660
gactcctgtt cacctggtat gactctctca ccggggttcc ggtcagccag cagaacctgc 720
tgctggagaa ggccagtgtc ctgttcaaca ctggggccct ctacacccag attgggaccc 780
ggtgtgatcg gcagacgcag gctgggctgg agagtgccat agatgccttt cagagagccg 840
caggggtttt aaattacctg aaagacacat ttacccatac tccaagttac gacatgagcc 900
ctgccatgct cagcgtgctc gtcaaaatga tgcttgcaca agcccaagaa agcgtgtttg 960
agaaaatcag ccttcctggg atccngaatg aattcttcat gctggtgaag gtggctcagg 1020
aggctgctaa ggtgggagag gtctaccaac agctacacgc agccatgagc caggcgccgg 1080
tgaaagagaa catcccctac tcctgggcca gcttagcctg cgtgaaggcc caccactacg 1140
cggccctggc ccactacttc actgccatcc tcctcatcga ccaccaggtg aagccaggca 1200
cggatctgga ccaccaggag aagtgcctgt cccagctcta cgaccacatg ccagaggggc 1260
tgacaccctt ggccacactg aagaatgatc agcagcgccg acagctgggg aagtcccact 1320
tgcgcagagc catggctcat cacgaggagt cggtgcggga ggcgagcctc tgcaagaagc 1380
tgcggacgat tgaggtgcta cagaaggtgc tgtgtgccgc acaggaacgc tcccggctca 1440
cgtacgccca gcaccaggag gaggatgacc tgctgaacct gatcgacgcc cccagtgttg 1500
ttgctaaaac tgagcaagag gttgacatta tattgcccca gttctccaag ctgacagtca 1560
cggacttctt ccagaagctg ggccccttat ctgtgttttc ggctaacaag cggtggacgc 1620
ctcctcgaag catccgcttc actgcagaag aaggggactt ggggttcacc ttgagaggga 1680
acgcccccgt tcaggttcac ttcctggatc cttactgctc tgcctcggtg gcaggagccc 1740
gggaaggaga ttatattgtc tccattcagc ttgtggattg taagtggctg acgctgagtg 1800
aggttatgaa gctgctgaag agctttggcg aggacgagat cgagatgaaa gtcgtgagcc 1860
tcctggactc cacatcatcc atgcataata agagtgccac atactccgtg ggaatgcaga 1920
aaacgtactc catgatctgc ttagccattg atgatgacga caaaactgat aaaaccaaga 1980
aaatctccaa gaagctttcc ttcctgagtt ggggcaccaa caagaacaga cagaagtcag 2040
ccagcacctt gtgcctccca tcggtcgggg ctgcacggcc tcaggtcaag aagaagctgc 2100
cctcccettt cagccttctc aactcagaca gttcttggta ctaatgtgag gaaacaaaca 2160
tgttcaggcc ccgaacattt ccggtgctga ctcggcctta aacgtttgtg ccataatgga 2220
aaatatctat ctatctgttc tcaaatcctg tttttctcat agtgtaaact cacatttgat 2280
gtgtttttat gaaggaaagt aaccaagaaa cctctaggaa ttagtgaaaa aagaactttt 2340
ttgaggtgtg ttactatact gctgtaagtt atttattata taaagtattg taaatagaat 2400
agtgttgaag atatgaaata tggctatttt taatggtgac aattatgact tttagtcact 2460
attaaattgg ggttacctat atcagtacaa tttgtagttg tttccaggtt tggctaataa 2520
tcattcctta acctagaatt cagatgatcc tggaattaag gcaggtcaga ggactgtaat 2580
gatagaatta aattagtgtc actaaaaact gtcccaaagt gctgcttcct aataggaatt 2640
cattaaccta aaacaagatg ttactattat atcgatagac tatgaatgct atttctagaa 2700
aaagtctagt gccaaatttg tcttattaaa taaaaacaat gtaggagcag cttttcttct 2760
agtttgatgt catttaagaa ttactaacac agtggcagtg ttagatgaag atgctgtcta 2820
caaggtagat aatatactgt ttgatactca aaacattttt cattttgttt aaagtagaag 2880

ttacataatt ctatatttta agtcttgggt aaaaaagtag ttttacattt tataaagtaa 2940
agatgtaaat gattcagctt taaagctcta tttgacttcc ttctttttgtc tgagatagcg 3000
tccagactgc gaaaagcga 3019

<210> 128
<211> 2312
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3576329CB1

<400> 128
gccggcgcgc ggtggggcat ggcgggttcg cggggtgcgg ggcgcacggc ggcgccgagc 60
gtgcggccgg agaagcggcg gtctgagccc gaactggagc ctgagcccga gccggagccc 120
cccctcctct gcacctctcc tctcagccac agcaccggca gcgattctgg cgtctccgac 180
agcgaggaga gtgtgttctc aggcctggaa gattccggca gtgacagcag tgaggatgat 240
gacgaaggcg acgaggaggg agaggacgga gcccttgatg acgagggcca cagtgggatt 300
aaaaagacca ctgaggagca ggtgcaggcc agcactcctt gcccgaggac agagatggcg 360
agcgcccgga ttggggatga gtatgcggag gacagctctg atgaggagga catccggaac 420
acggtgggca acgtgccctt ggagtggtac gatgacttcc cccacgtggg ctacgacctg 480
gatggcaggc gcatctacaa gcccctgcgg acccgggatg agctggacca gttcctggac 540
aagatggacg atcctgacta ctggcgcacc gtgcaggacc cgatgacagg gcgggacctg 600
agactgacgg atgagcaggt ggccctggtg cggcggctgc agagtggcca gtttggggat 660
gtgggcttca acccctatga gccggctgtc gacttcttca gcggggacgt catgatccac 720
ccggtgacca accgcccggc cgacaagcgc agcttcatcc cctccctggt ggagaaggag 780
aaggtctctc gcatggtgca cgccatcaag atgggctgga tccagcctcg ccggccccga 840
gaccccaccc ccagcttcta tgacctgtgg gcccaggagg accccaacgc cgtgctcggg 900
cgccacaaga tgcacgtacc tgctcccaag ctggccctgc caggccacgc cgagtcgtac 960
aacccaccccc ctgaatacct gctcagcgag gaggagcgct tggcgtggga acagcaggag 1020
ccaggcgaga ggaagctggg ctttttgcca cgcaagttcc cgagcctgcg ggccgtgcct 1080
gcctacggac gcttcatcca ggaacgcttc gagcgctgcc ttgacctgta cctgtgccca 1140
cggcagcgca agatgagggt gaatgtagac cctgaggacc tcatccccaa gctgcctcgg 1200
ccgagggacc tgcagccctt ccccacgtgc caggccctgg tctacagggg ccacagtgac 1260
cttgtccggt gcctcagtgt ctctcctggg ggccagtggc tggtttcagg ctctgacgac 1320
ggctccctgc ggctctggga ggtggccact gcccgctgtg tgaggactgt tcccgtgggg 1380
ggcgtggtga agagtgtggc ctggaacccc agccccgctg tctgcctggt ggctgcagcc 1440
gtggaggact cggtgctgct gctgaaccca gctctggggg accggctggt ggcgggcagc 1500
acagatcagc tgttgagcgc cttcgtcccg cctgaggagc ccccccttgca gccggcccgc 1560
tggctggagg cctcagagga ggagcgccaa gtgggcctgc ggctgcgcat ctgccacggg 1620
aagccagtga cgcaggtgac ctggcacggg cgtggggact acctggccgt ggtgctggcc 1680
acccaaggcc acacccaggt gctgattcac cagctgagcc gtcgccgcag ccagagtccg 1740
ttccgccgca gccacggaca ggtgcagcga gtggccttcc accctgcccg gcccttcctg 1800
ttggtggcgt cccagcgcag cgtccgcctc taccacctgc tgcgccagga gctcaccaag 1860
aagctgatgc ccaactgcaa gtgggtgtcc agcctggcgg tgcaccctgc aggtgacaac 1920
gtcatctgtg ggagctacga tagcaagctg gtgtggtttg acctggatct ttccaccaag 1980
ccatacagga tgctgagaca ccacaagaag gctctgcggg ctgtggcctt ccacccgcgg 2040
tacccactct ttgcgtcagg ctcggacgac ggcagtgtca tcgtctgcca tggcatggtg 2100
tacaatgacc ttctgcagaa ccccttgctg gtgcccgtca aggtgctgaa gggacacgtg 2160
ctgacccgag atctgggagt gctggacgtc atcttccacc ccacccagcc gtgggtcttc 2220
tcctcggggg cagacgggac tgtccgcctc ttcacctagc tgttctgcct gcctggggct 2280
ggggtggtcg tgctgaagtc aacagagcct tc 2312

<210> 129
<211> 921
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 3805550CB1

<400> 129
aggcggagtc ggggcggtgt gctgaggtgg gcctgagggc ggagtcgagg tcgggctgaa 60
ggcggagtcg gggagggctg aggtgggcct gaaggcagag tcgaggccat ggcagggccg 120
ggcccaggcc cgggggaccc ggacgagcag tacgatttcc tgttcaagct ggtgctggtg 180
ggcgacgcaa gcgtgggcaa gacgtgcgtg gtgcagcgct tcaagaccgg cgccttctcg 240
gagcgccagg gaagcaccat cggcgtcgac ttcaccatga agacgctgga gatccagggc 300
aagcgggtca agctgcagat ctgggacacg gccggccagg agcggttccg caccatcacc 360
cagagctact accgcagtgc caatggggcc atccttgcct acgacatcac caagaggagc 420
tccttcctgt cggtgcctca ctggattgag gatgtgagga agtatgcggg ctccaacatt 480
gtgcagctgc tgatcgggaa caagtcagac ctcagcgagc ttcgggaggt ctccttggct 540
gaggcacaga gcctggctga gcactatgac atcctgtgtg ccattgagac gtctgccaag 600
gactcgagca acgtggagga ggccttcctg agggtggcca cggagctcat catgcggcac 660
gggggcccct tgttcagcga gaagagcccc gaccacatcc agctgaacag caaggacatc 720
ggagaaggct ggggctgcgg gtgctgacca ggggccgggc cggcagactg gggttcccc 780
acctccttgc tctccccagc ctgccaagcc cagccctcca gagccagccc tcctgggtac 840
cggcaactac agcagccggg tgaagctctg gagctctgca tcctgtggcc tggctgcggg 900
atggaggctc tccttgagga a 921

<210> 130
<211> 1291
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4546403CB1

<400> 130
ctcgagcgaa tcggctcgag agatggctcc ttggcggcat gtgcattttc tcctaatgga 60
agcttctttg tcactggcac ttcatgtggt gatttaacag tgtgggatga tcaaatgagg 120
tgtctgcata gtgaaaaagc acatgatctt ggaattacct gctgcgattt ttcttcacag 180
ccagtttctg atggagaaca aggtcttcag ttttttcgac tggcatcatg tggtcaggat 240
tgccaagtca aaatttggat tgtttctttt acccatatct taggttttga attaaaatat 300
aaaagtacac tgagtgggca ctgtgctcct gttctggctt gtgcttttcc ccatgatggg 360
cagatgctag tctcagggtc agtggataag tctgtcatag tatatgatac taatactgag 420
aatatacttc acacattgac tcagcacacc aggtatgtca caacttgtgc ttttgcacct 480
aataccccttt tacttgctac tggttcaatg gacaaaacag tgaacatctg gcaatttgac 540
ctggaaacac tttgccaagc aaggagcaca gaacatcagc tgaagcaatt taccgaagat 600
tggtcagagg aggatgtctc aacatggctt tgtgcacaag atttaaaaga tcttgttggt 660
attttcaaga tgaataacat tgatggaaaa gaactgttga atcttacaaa agaaagtctg 720
gctgatgatt tgaaaattga atctctagga ctgcgtagta aagtgctgag gaaaattgaa 780
gagctcagga ccaaggttaa atccctttct tcaggaattc ctgatgaatt tatatgtcca 840
ataactagag aacttatgaa agatccggtc atcgcatcag atggctattc atatgaaaag 900
gaagcaatgg aaaattggat cagcaaaaag aaacgtacaa gtcccatgac aaatcttgtt 960
cttccttcag cggtacttac accaaatagg actctgaaaa tggccatcaa tagatggctg 1020
gagacacacc aaaagtaaaa ttgttgatat tgtattattt atattttcag tgatctcatt 1080
tgaatgattt ataggtaaat actaatcaga cattattaaa agcaaaacag gaaaaaggta 1140
aacttcttaa atttagttac ctataaaaat tgtcaatttt cattctttaa aaacacatgg 1200
acttactata aaagcctttt tgtactagtg aaaagaatct tcagctatat agaaataaag 1260
ttatacttta aattgcaaaa aaaaaaaaaa a 1291

<210> 131
<211> 1836
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4767318CB1

<400> 131
ttttagaagg ttagtgttgg ttcttttatt cgattaaaca ggaatacaca tatgtctacc 60
aaagaatagg taagggagaa ataagaacac taaaaaaact cggaatcgtt aagtgtgaag 120

catatttgga gttaaaagaa ccaaatatta ctaagtaagc agacgcgggc acgcgctgca 180
taccgggatt tgtagtccct tccggggcgg ggtacagcgc gcctgcgcag aggggccgtc 240
gctcttccgg gcgcatgcgt gcggcagcgg cgccaggact gactgcgccg tggaggctgc 300
tgcagtgttg tgagttggaa gctggggagc tcggcatggc ggtccccgct gcagccatgg 360
ggccctcggc gttgggccag agcggccccg gctcgatggc cccgtggtgc tcagtgagca 420
gcggcccgtc gcgctacgtg cttgggatgc aggagctgtt ccggggccac agcaagacgc 480
gcgagttcct ggcgcacagc gccaaggtgc actcggtggc ctggagttgc gacgggcgtc 540
gcctagcctc ggggtccttc gacaagacgg ccagcgtctt cttgctggag aaggaccggt 600
tggtcaaaga aaacaattat cggggacatg gggatagtgt ggaccagctt tgttggcatc 660
caagtaatcc tgacctattt gttacggcgt ctggagataa aaccattcgc atctgggatg 720
tgaggactac aaaatgcatt gccactgtga acactaaagg ggagaacatt aatatctgct 780
ggagtcctga tgggcagacc attgctgtag gcaacaagga tgatgtggtg acctttattg 840
atgccaagac acaccgttcc aaagcagaag agcagttcaa gttcgaggtc aacgaaatct 900
cctggaacaa tgacaataat atgttcttcc tgacaaatgg caatggttgt atcaacatcc 960
tcagctaccc agaactgaag cctgtgcagt ccatcaacgc ccatccttcc aactgcatct 1020
gtatcaagtt tgaccccatg ggaagtactt ttgccacagg aagtgcggat gctttggtca 1080
gcctctggga tgtggatgag ttagtgtgtg ttcggtgctt ttccaggctg gattggcctg 1140
taagaaccct cagtttcagc catgatggga aaatgctggc gtcagcatcg gaagatcatt 1200
ttattgacat tgctgaagtg gagacagggg acaaactatg ggaggtacag tgtgagtctc 1260
cgaccttcac agtggcgtgg caccccaaaa ggcctctgct ggcatttgcc tgtgatgaca 1320
aagacggcaa atatgacagc agccgggaag ccggaactgt gaagctgttt gggcttccta 1380
atgattcttg agaggaggtt gtagggagag gaggccccgg cagaggtctt ccttcatgtg 1440
gttagtttgg tctgttctct cggagttggt gggcacccta aatatttgta agttggtata 1500
aattgtaaac gtctctggtc aggctgcgca tttcgttctt ttgctttgtc tgtgtattag 1560
ctctttccat tctttgcccc cagcatgagt taactcgcgt ggactctgca gtgcgagtag 1620
tgaccccagc ataccttgtc ctctggacct cctgtcttct ctgcttctgg gtgcatggta 1680
gactttgtgg catttgatac aacttggaca atacctagtt tggagggagg ggaatggaag 1740
ggcatggaag ttttttttaaa taattaaaaa tatatacata taattttgag aattgagcat 1800
ttaataaact gacttttgtt attatggaaa aaaaaa 1836

<210> 132
<211> 2136
<212> DNA
<213> Homo sapiens

<220>
<221> misc_feature
<223> Incyte ID No: 4834527CB1

<400> 132
ggcgcgccgg gagccggcag acatgccaca gacgctgagt gcctccgaca tggtcacccc 60
aggcagcctc agcccacccc ccaccgagcc cacagatggc gaacaggctg ggcagcccct 120
cctggatgga gcgccatcct cagcctccct ggaaacactg atccagcacc tggtgcccac 180
agccgactac taccccgaga aagcctacat cttcaccttc ctgctgagct ctcgcctctt 240
catcgagccc cgggagctcc tggcccgggt ctgccacctg tgcatcgagc agcagcagct 300
ggacaagccg gtgctggaca aggcccgggt ccggaagttc ggccccaaac tgctgcagct 360
gttggccgag tggaccgaga ccttcccaag ggacttccag gaagagtcga ctatcgggca 420
ccttaaggac gtcgtgggcc gcatcgcccc ctgtgacgag gcataccgga agaggatgca 480
tcagctccta caggctctgc accagaagct ggcggctctg cgccaggggc cagaaggtct 540
ggtgggtgcc gacaagccca tctcctacag gaccaagcca ccagcctcca tccacaggga 600
gctccttggt gtctgcagcg acccctacac actggcccag cagctgaccc acgtggaact 660
ggagcggctg cggcacatcg ggcctgagga gtttgtccag gcctttgtga acaaggaccc 720
tctggccagc acaaagccct gcttcagtga caagaccagc aacctggagg cttatgtgaa 780
atggttcaac aggctgtgct acctggtggc aactgagatc tgcatgccag ccaagaagaa 840
gcagagggcc caggtgattg agttcttcat cgacgtggcc cgcgagtgct tcaacatcgg 900
caacttcaac tccctcatgg ccatcatctc cggcatgaac atgagccctg tctccaggct 960
gaagaagacc tgggccaaag tgaggacggc caagtttttc atcctcgagc accagatgga 1020
cccaacgggg aatttctgca actacaggac agccctgcgc ggggcggccc accgctccct 1080
gacggcccac agcagccgag agaagattgt cattcctttc ttcagcctgc tcatcaaaga 1140
catctacttc ctgaatgagg gctgcgccaa ccgccttccc aatggacacg tcaactttga 1200
gaaattcctg gagctggcca agcaggtggg ggagttcatc acctggaaac aagtggagtg 1260
tcccttcgag caagacgcca gcatcaccca ctacctgtac accgccccca tcttcagtga 1320
ggatggtctt tatttggctt cttatgaaag tgagagccca gagaaccaaa cagaaaaaga 1380

```
aagatggaaa gctctaagat cttctatttt ggggaagaca tgaaagcgct gagctgaggg 1440
acgaggaaga gctggagccc gcagaagccg tccacagccc tgcctcagtg gcccagtggg 1500
cagaggccag ggagtgcctc actattttgc aaatgccgac cctgtggcct gctgcccgcc 1560
ccccgccccc cacagtggcc atacgggcac aggagacctt ttatgggact ttggccctgg 1620
caggacccag ggcctccaga cgtgcgggcg gcacatgcct tggggacatc ctgccttcag 1680
gaccgtgggg cctggtcagt ctgtccatcc tcggcaagga cacaacactg ccccagaggg 1740
tgggaccact gcaagctcga gaccttgctt ggtgacatgt gccactttgg ccaccaccca 1800
cagtctgtca ccacgtggct tgggaacttc tggagccaca gcaggcatca cggtgcgacg 1860
tgagatgcct gcgccagccc cgagcccact ggcagccact gccattccac ccatggtccc 1920
tcaccctgcc ctgccgacga gcttgtctct gcagccccag gtaccccctt cctggatgct 1980
gctggcccca ggagatagct ttccggtgac agctgtggaa cgcgtcagca ggacaaactg 2040
gacacatgga gttacagtgt gtacacggca gtcccgccac ccagcccccct tgtaaactct 2100
agtcactata aacacacccg tacgcctaaa aaaaaa                          2136
```